#FranceIA

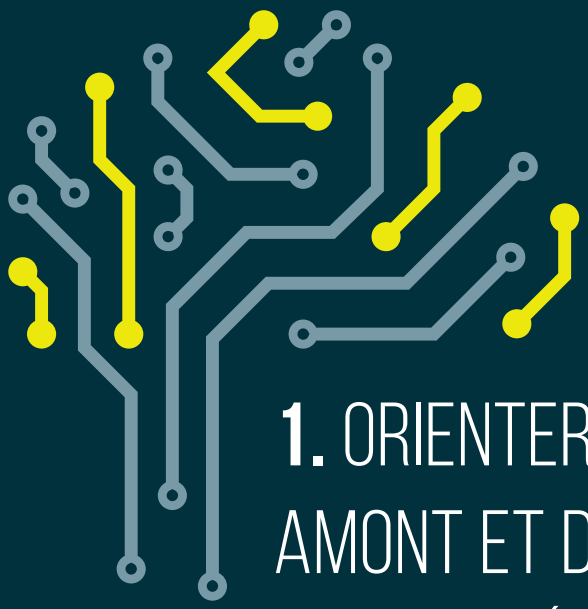
FRANCE
INTELLIGENCE
ARTIFICIELLE

RAPPORT DE SYNTHÈSE
FRANCE INTELLIGENCE ARTIFICIELLE

GROUPES DE TRAVAIL

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
Le Gouvernement

Sommaire



1. ORIENTER LA RECHERCHE AMONT ET DÉVELOPPER LES COMPÉTENCES

GROUPE DE TRAVAIL 1.1
RECHERCHE AMONT

Groupe Recherche Amont

Périmètre d'étude du groupe de travail

Les buts de ce groupe de travail sont de:
1. Identifier les grands domaines de recherche fondamentale
2. Identifier les sujets disruptifs de recherche (impact potentiel et applications)
3. Identifier l'expertise disponible en France sur les sujets disruptifs
4. Identifier les moyens à mettre en œuvre (partenariats nationaux, coopérations internationales, financements …)
5. Prioriser les sujets disruptifs
6. Etablir des recommandations concernant la recherche amont

Etat des lieux

La recherche en intelligence artificielle (en France et à l'international) est morcelée en un grand nombre de sous-communautés, ayant pour certaines une très grande autonomie. Cela est du à la fois au grand nombre de questions différentes et aux différents outils formels utilisés. Nous avons fait un état des lieux des grands domaines de recherche, que nous pouvons catégoriser ainsi:
(voir en Annexe 4 la description de chacun de ces grands domaines de recherche)

- Apprentissage automatique (machine learning), en particulier deep learning
- Représentation des connaissances et modélisation des raisonnements
- Décision et gestion de l'incertitude
- Satisfaction de contraintes et SAT (satisfaisabilité d'une formule logique)
- Planification et recherche heuristique
- Agents autonomes et systèmes multi-agents
- IA & web
- Traitement automatique des langues
- Robotique
- Vision et reconnaissance des formes
- Modélisation cognitive et systèmes neuro-informationnels

Il faut également noter la proximité de certaines thématiques développées en intelligence artificielle avec d'autres domaines de l'informatique, en particulier l'informatique théorique, la recherche opérationnelle, les bases de données (et en particulier le big data), et l'interaction humain-machine.

Et également les liens existants et potentiels avec d'autres disciplines, notamment les mathématiques (en particulier les statistiques), la philosophie, la linguistique, l'économie, les sciences cognitives, et les neurosciences.

Il nous semble important de souligner cette diversité thématique, car il est nécessaire de réaliser que l' « intelligence artificielle » n'est pas un tout uniforme, et de la prendre en compte pour les recommandations et décisions qui seront prises.

En particulier les applications qui ont été visibles récemment (Alphago, Watson, Libratus, etc.) sont des applications nécessitant l'intégration de plusieurs notions et techniques issues de thématiques différentes. Il y a donc à la fois un besoin de développer les notions et techniques dans chacune des thématiques, mais également de favoriser l'intégration de ces notions et techniques pour attaquer des problèmes réels importants.

Le groupe de travail veut souligner les grandes potentialités françaises en intelligence artificielle, la France ayant des chercheurs du meilleur niveau international dans chacune de ces thématiques (ce qui n'est pas le cas de tous les pays), et ce qui ne limite donc pas le champ des possibles.

Le groupe de travail constate également le manque de grands groupes industriels dans la recherche amont en intelligence artificielle en France, contrairement à ce qui peut se faire dans d'autres pays.

IDENTIFICATION DES SUJETS DE RECHERCHE DISRUPTIFS

Au niveau des grandes questions de recherche, nous avons identifié 36 sujets de recherche disruptifs. Nous avons classés ces sujets en 9 grandes thématiques, que nous détaillons ci-dessous:
(voir en Annexe 3 la description de chacun de ces sujets de recherche disruptifs)

- Perception
La perception et les informations perceptuelles sont une entrée de beaucoup de systèmes de raisonnement. Par exemple, pour concevoir des robots capables d'interagir avec leur environnement les capacités de perception sont le premier chainon indispensable avant de mener un raisonnement, de prendre une décision et d'agir. Or beaucoup reste à faire sur ces problématiques liées à la perception.
Les grands défis concernant cette thématique:
- Vision
- Parole / Audio
- Multimodalité

- Collaboration Humain Machine
Cette problématique est primordiale pour améliorer les interactions humains-machines, afin de permettre aux utilisateurs d'avoir une meilleure interaction avec les programmes et robots. Cela passe en particulier par la facilitation de l'interaction et donc le développement d'outils pour le traitement du langage naturel.
Les grands défis concernant cette thématique:
- Recommandations à l'utilisateur
- Décision sous incertitude
- Sémantique et compréhension de la langue
- Intentionnalité / Emotions

- Intelligence Artificielle à large échelle
Un des éléments saillants à l'heure actuelle est le fait que dans de nombreux domaines une grande quantité de données est disponible, et il faut adapter les méthodes usuelles (raisonnement, décision, apprentissage, etc) pour pouvoir appréhender efficacement ces données.
Les grands défis concernant cette thématique:
- Raisonner sur des données massives
- Intégration de données hétérogènes et/ou multisources
- Aspects matériels et calcul (architectures distribuées, etc.)

- Explicabilité
Une capacité importante pour les systèmes artificiels automatisés est le fait de pouvoir expliquer, argumenter, justifier un raisonnement, une décision ou une action. Cela est indispensable pour une meilleure acceptabilité des programmes prodiguant conseils ou injonctions (dans le cas d'un véhicule autonome par exemple).
Les grands défis concernant cette thématique:
- Résumé / Restitution linguistique
- Argumentation / Délibération
- Explicabilité des méthodes d'apprentissage
- Causalité

- Apprentissage machine

La puissance des techniques d'apprentissage est illustrée par le déploiement de celles-ci pour un nombre impressionnant d'applications (notamment la vision, la traduction automatique, les véhicules autonomes, etc.). Ces techniques sont applicables principalement lorsque l'on dispose de masses de données. Mais il reste à travailler sur un grand nombre de points, que ce soit sur les fondements de ces approches, ou sur les cas où l'on dispose de moins d'exemples.

Les grands défis concernant cette thématique:

- Apprentissage profond
- Apprentissage à partir de peu d'exemples / non supervisé
- Apprentissage en ligne et incrémental
- Apprentissage par renforcement
- Transfert / Analogie

- Résolution de Problèmes

Le *machine learning* permet, mieux que jamais, d'aborder la résolution de problèmes qui nous semblent intuitifs (reconnaître des entités sur des images) au travers de grandes quantités d'exemple s annotés. Mais une IA doit aussi être capable de résoudre des problèmes de raisonnement logique bien posés, difficiles pour l'être humain comme pour la machine, qu'il s'agisse de prouver des théorèmes ou de résoudre des casse-têtes de tout ordre, y compris ceux d'intérêt pratique. . .

Les grands défis concernant cette thématique:

- Problèmes NP-difficiles (Beyond NP)
- Optimisation
- Recherche heuristique et méta-heuristique
- Raisonnement Multi-Echelle

- Décision, agents autonomes et intelligence collective

Il y a beaucoup d'applications potentielles demandant l'intervention d'agents autonomes intégrant des capacités de perception, d'apprentissage, raisonnement, décision et d'interaction, ou de plusieurs agents, qui peuvent être aussi bien des entités logicielles (sur internet) ou matérielles (robots). On peut imaginer par exemple le développement d'assistants personnels représentant et négociant pour leur propriétaire, ou des applications où plusieurs robots doivent collaborer pour le secours de personnes, etc. Il est crucial de disposer des outils formels nécessaires à la modélisation et à la régulation de ces interactions.

Les grands défis concernant cette thématique:

- Choix social computationnel / Théorie des jeux
- Raisonnement sur l'autre
- Actions conjointes / Planification collaborative
- IA intégrative et robotique

- Intelligence artificielle générale

Au delà du développement nécessaire de techniques dédiées à des problèmes particuliers, il est important de développer des techniques permettant de s'adapter à des situations peu ou pas connues à l'avance. Ce que nous avons résumé en « intelligence artificielle générale ».

Les grands défis concernant cette thématique:

- Informatique neuro-inspirée
- IA embarquée
- Créativité et génération
- Jeux génériques (General Game Playing)
- Intégration de différents modes de raisonnement

- Enjeux éthique et sociaux, acceptabilité, et protection de la vie privée

Il sera nécessaire de travailler sur les enjeux éthiques et sociaux des méthodes d'intelligence artificielle. Si une grande partie de ces considérations concerne directement l'éthique et les sciences sociales, et sort donc du cadre de ce groupe sur la recherche amont en IA, certaines de ces questions sont des sujets de recherche à part entière en IA.

Les grands défis concernant cette thématique:

- Intégration de la préservation de la vie privée dans les méthodes d'apprentissage et de raisonnement
- Acceptabilité et appropriation
- Aspects déontiques (respect des lois par les méthodes d'intelligence artificielle)
- IA pour le droit (modélisation du raisonnement juridique)

Enjeux

L'enjeu principal est d'affirmer le leadership mondial de la France sur la recherche en intelligence artificielle.

Les enjeux plus précis que nous avons identifiés pour atteindre ce but sont:

- **Renforcer le soutien sur tout le front des recherches thématiques**

Les avancées de la recherche amont d'aujourd'hui sont les moteurs potentiels d'innovation de demain. Il faut avoir à l'esprit que les principes de l'apprentissage profond, qui révolutionnent beaucoup de domaines actuellement, ont été inventés il y a 25 ans, et qu'il a fallu plus de 10 ans, et l'évolution vers des conditions favorables (puissance de calcul, disponibilité des données, etc) pour faire éclore cette potentialité. Il est donc indispensable de permettre aux chercheurs en intelligence artificielle de développer leurs travaux sereinement, afin de découvrir les techniques qui permettront les prochaines innovations.

- **Favoriser la recherche aux interfaces**

La plupart des applications actuelles et potentielles de l'intelligence artificielle nécessitent d'utiliser conjointement plusieurs techniques issues de domaines différents. Cela s'illustre facilement avec les applications qui ont fait parler d'elles ces derniers mois (alphago, watson, libratus, etc.). Mais cette intégration pose des difficultés propres et nécessite des incitations dédiées, afin de décloisonner les différents domaines. Il s'agit de faciliter d'une part les liens entre domaines de l'intelligence artificielle (étudier par exemple les liens entre approches symboliques et approches numériques, apprentissage et raisonnement, ou apprentissage et résolution de contraintes), et d'autre part entre l'intelligence artificielle et d'autres disciplines (mathématiques, économie, philosophie, linguistique, droit, SHS, physique, biologie, médecine, etc.)

- **Maintenir et renforcer les forces de travail en France**

On constate à l'heure actuelle une fuite d'un certain nombre de nos collègues brillants vers des entreprises privées internationales. Si on peut se réjouir du fait que cela montre la reconnaissance du niveau de la formation et de la recherche française en intelligence artificielle, il y a un risque d'appauvrissement des compétences nationales. La question est de savoir quoi faire pour tenter de retenir nos collègues et leur expertise, et d'éviter les problèmes organisationnels liés à ces départs. Une problématique associée serait de trouver les moyens de renforcer l'attractivité internationale.

- **Rapprocher les industriels de la recherche amont**

Un moteur important d'intégration est le fait de tenter de résoudre un problème particulier. Un enjeu potentiel dans le cadre du PIA pourrait être de favoriser cette recherche intégrative, en intégrant plusieurs grands programmes, avec le concours d'industriels amenant des besoins en applications innovantes (par le biais de challenges par exemple), afin de mobiliser une partie de la communauté sur ces collaborations entre domaines. Dans d'autres pays, les industriels financent la recherche amont qui les intéresse directement, sans incitation étatique.

- **Accès et préparation des données - Disponibilité des codes et plate-formes**

Il semble nécessaire d'avoir une initiative nationale pour créer des corpus exploitables et des bases de benchmarks. Cela devrait être accompagné d'une réflexion sur les annotations, les biais, et les problèmes de vie privée. Plus généralement la création d'une plate-forme de partage de jeux de données

et de code de la recherche permettrait à la fois de régler des problèmes d'accessibilité aux codes et aux données, afin d'accélérer les travaux des uns et des autres, mais également d'améliorer les pratiques scientifiques, notamment en termes de reproductibilité.

- Favoriser la prise de risque, fluidifier et simplifier l'accès aux financements
Dans un domaine aussi concurrentiel et dynamique que l'intelligence artificielle, les lourdeurs administratives liées à l'évaluation et au financement de la recherche peuvent être un frein important. Un des enjeux cruciaux pour permettre d'accroitre les performances dans le domaine serait de favoriser la prise de risque (il n'est pas possible, avec les financements actuels, de se lancer dans des recherches très prospectives, sans « livrables » et publications à court terme), ainsi que de fluidifier et de simplifier l'accès aux financements, pour qu'un chercheur avec une idée nouvelle n'ait pas à attendre typiquement 1 an (attendre le prochain appel à projet, puis le temps de l'évaluation) pour avoir les moyens de la mettre en oeuvre. Cela semble nécessaire pour ne pas freiner l'émergence de nouveaux sujets disruptifs.

- Accès à des ressources matérielles de calcul dédiées
Les progrès de certaines thématiques dépendent beaucoup des ressources de calcul disponibles, afin de pouvoir tester expérimentalement la validité et l'efficacité des approches. Or un certain nombre de collègues n'ont pas la chance d'avoir dans leur laboratoire les ressources matérielles adéquates. C'est particulièrement vrai pour les personnes travaillant ou utilisant l'apprentissage, qui nécessite des architectures dédiées (GPU). Ce qui est donc à frein à leur recherche.

Recommandations

Voici les recommandations du groupe de travail recherche amont, étant donné les sept enjeux identifiés à la section précédente.
L'ensemble des coûts indiqués pour ces recommandations sont calculés pour un horizon sur 10 ans.

Projets Initiative Intelligence Artificielle (initIA)

Enjeux concernés:

- Renforcer le soutien sur tout le front des recherches thématiques
- Maintenir et renforcer les forces de travail en France
- Favoriser la prise de risque, fluidifier et simplifier l'accès aux financements

Paragraphe de synthèse:

Mise en place d'un financement attribué à un seul porteur et son équipe proposant un projet innovant et ambitieux (avec prise de risque) relevant de l'IA (et pouvant inclure un certain nombre de collègues), sur une longue période (typiquement 5 ans), avec deux niveaux de financements/difficultés (typiquement jeune chercheur / chercheur confirmé). Ces financements doivent être suffisants pour pouvoir mener un projet de recherche, et pour financer des thèses et/ou post-doctorants (typiquement au moins 500k€ pour les jeunes chercheurs et 1M€ pour les chercheurs confirmés). Le taux de sélection doit être suffisamment attractif. Nous recommandons 40 bourses chercheurs confirmés et 30 bourses jeunes chercheurs par an pendant 10 ans. Le dossier et la procédure de candidature devront être légers (de l'ordre de quelques pages). L'obtention d'un tel financement pourrait être conditionné par l'engagement de candidater à un projet européen (ERC ou autre) avant la fin du financement.

Commentaires:

Nous insistons sur l'aspect stratégique pour le maintien des forces de travail en IA en France de cette recommandation: une part importante de nos collègues (jeunes et seniors) quittent actuellement l'académie pour le privé (souvent dans des entreprises étrangères). Il n'est pas possible de concurrencer le privé au niveau des salaires, mais une partie non négligeable de cette décision est également motivée par des conditions de travail bien meilleures. Avoir ce type de financement pour permettre aux meilleurs de développer leur recherche dans de bonnes conditions permettrait d'avoir un réel levier pour la recherche académique française.
Nous insistons également sur l'aspect complémentaire de ces financements ciblés au niveau des équipes, avec les financements traditionnels qui se tournent plus sur l'organisation en réseau (par exemple dans le cas des ANR). Il serait à ce titre également de première importance que l'ANR lance des appels dédiés à l'intelligence artificielle. Il serait également souhaitable que « la conception d'agents autonomes » soit ajoutée à la liste des défis de l'ANR.
Au niveau des thématiques visées pour ces projets, nous préconisons de focaliser les appels sur les sujets disruptifs que nous avons identifiés (voir ci-dessus et l'Annexe 3), et organisés en 9 grandes thématiques: perception; collaboration humain machine; intelligence artificielle à large échelle; explicabilité; apprentissage machine; résolution de problèmes; décision, agents autonomes et intelligence collective; intelligence artificielle générale; enjeux éthique et sociaux, acceptabilité, et protection de la vie privée.
La France ne soumet pas suffisamment de candidatures ERC, ces financements peuvent donc également être considérés comme des moyens d'augmenter le nombre de candidatures (et donc de succès espérés) à l'ERC (et aux autres programmes européens), afin d'assurer le leadership de la France au niveau de l'intelligence artificielle européenne.

Coût: 550M€

Projets aux interfaces (IA+X)

Enjeux concernés:

- Favoriser la recherche aux interfaces

Paragraphe de synthèse:

Soutenir des projets collaboratifs nécessitant une recherche aux interfaces. Il s'agit d'une part des interfaces au sein de l'IA, pour favoriser la recherche intégrative. Et d'autre part les interfaces entre IA et un autre domaine (mathématiques, économie, philosophie, linguistique, droit, SHS, physique, biologie, médecine, etc.). Ces projets devraient être typiquement d'une durée de 5 ans. Ils devront être portés par au moins 2 chercheurs principaux de domaines différents, dont au moins un d'IA.

Coût: 300M€

Très grande infrastructure pour l'Intelligence Artificielle

Enjeux concernés:

- Accès et préparation des données - Disponibilité des codes et plate-formes
- Accès à des ressources matérielles de calcul dédiées

Paragraphe de synthèse:

Lancer une très grande infrastructure pour l'IA. Ce qui inclut:
1) des capacités de calcul (en particulier GPU) pour les équipes d'intelligence artificielle.

2) la mise en place de bases de données, de données étiquetées, et de corpus pour accélérer les travaux en IA.
3) Il faudra veiller à la mise en place de corpus en Francais. 4) une plate-forme d'accès aux logiciels développés par les différentes équipes (accès pour les autres équipes académiques, mais également pour les industriels, et les citoyens).
4) des personnels pour opérer ces plates-formes (logiciels, interface, et ASR).

Cette infrastructure est nécessaire pour permettre une accélération des développements en IA, ainsi que l'amélioration des pratiques scientifiques, notamment en termes de reproductibilité.
Au niveau organisationnel, soit une grande entité centrale regroupant les équipements les plus importants (en particulier la plate-forme logicielle et l'accès aux données), et pilotant d'autres structures décentralisées et plus proches des utilisateurs (pour les capacités de calcul), soit une construction en réseau.

Commentaires:

Au delà de fournir des moyens de calcul et de stockage (pour les jeux de données, en français pour le texte et la parole), le but est d'aller dans le même sens que ce qui se fait dans les plateformes de calcul en bioinformatique où l'on a accès aux logiciels (français ou pas, compilés, installés, utilisables). Ce type de service (calcul+données+logiciels installés) pourrait être particulièrement attirant pour les scientifiques mais aussi les IA-entrepreneurs (qui voudraient pouvoir facilement jouer avec un outil, sans investir sur le calcul, en capitalisant les problèmes d'installation et de mise à jour). Cela nécessite des forces humaines, dans la durée.

Coût: 150M€

Centre pour l'Intelligence Artificielle

Enjeux concernés:

- Renforcer le soutien sur tout le front des recherches thématiques
- Favoriser la recherche aux interfaces
- Maintenir et renforcer les forces de travail en France
- Rapprocher les industriels de la recherche amont

Paragraphe de synthèse:

Création d'un Centre pour l'Intelligence Artificielle Français, pour organiser des séminaires spécialisés (en IA ou pour l'interdisciplinarité), comme lieu de rencontre entre académiques et industriels (grands groupes et startups), pour héberger des équipes souhaitant lancer un projet commun (par exemple en vue de préparer une candidature à un programme européen), permettant également quelques résidences (financement et hébergement de chercheurs sur quelques mois pour travailler à un sujet donné). Au niveau organisationnel, cela nécessite donc un lieu avec amphithéatre et salles de réunion, capacité d'hébergement et de restauration. Le budget indiqué ne prend en compte que les aspects fonctionnement.

Commentaires:

Ici nous n'avons indiqué que les coûts liés au fonctionnement d'une telle structure. Il faudrait ajouter le coût du foncier correspondant si celle-ci bénéficie d'un nouveau bâtiment dédié, ou le coup de l'aménagement si la décision est d'utiliser du foncier déjà existant.
Ce centre serait à la fois un bon outil de création et de renforcement de collaborations, et d'accélération des échanges (entre académiques, et entre académiques et acteurs industriels), mais également un bel étendard national et international.

L'idée est de prendre exemple sur de tels centres qui existent en France dans d'autres disciplines, ou au niveau international en informatique (citons par exemple le Schloss Dagsthul, l'IHP, l'IHES, le Turing institute, l'IMA (Minéapolis)), et qui montrent leur efficacité à la fois scientifique et en termes de reconnaissance internationale.
Ce centre pourrait bénéficier à la fois de financements publics et de mécénats de grands acteurs industriels, pour souligner sa vocation à la fois académique et de transfert.

Coût: 70M€

Associer les industriels à la recherche amont

Enjeux concernés:

- Rapprocher les industriels de la recherche amont
- Maintenir et renforcer les forces de travail en France

Paragraphe de synthèse:

Cela correspond à 3 initiatives:
1) Cibler une partie des financements accordés aux industriels (que ce soit dans le cadre du CIR ou du PIA) sur des collaborations réelles et formalisées avec des équipes de recherche académiques.
2) Mettre en place des chaires industrielles dans les universités et les organismes de recherche (ces chaires pourraient être cofinancées par des industriels et des fonds du PIA par exemple).
3) Faciliter le temps partagé industrie/académique, en particulier en assouplissant les conditions administratives et financières de cumul pour les chercheurs titulaires.

Ces 3 initiatives correspondent à du ciblage de crédits déjà existants, ou à des modifications organisationnelles, elles n'engendrent donc pas de coûts propres.

Coût: 0M€

Annexes

Annexe 1 – Composition du groupe Recherche Amont

Le GT Recherche Amont est animé par Sébastien Konieczny (*Directeur de Recherche CNRS*, CRIL - Lens), et il est composé de:

- Jamal Atif, *Professeur*, LAMSADE - Paris
- Francis Bach, *Directeur de Recherche INRIA*, DI-ENS - Paris
- Claude Berrou, *Professeur émérite*, IMT Atlantique
- Isabelle Bloch, *Professeur*, LTCI Télécom ParisTech
- Raja Chatila, *Professeur*, ISIR - Paris
- Matthieu Cord, *Professeur*, LIP6 - Paris
- Béatrice Daille, *Professeur*, LS2N - Nantes
- Sébastien Konieczny, *Directeur de Recherche CNRS*, CRIL - Lens
- Jérôme Lang, *Directeur de Recherche CNRS*, LAMSADE - Paris
- Francois Pachet, Directeur de *Sony-CSL*, Sony-CSL - Paris
- Catherine Pelachaud, *Directeur de Recherche CNRS*, ISIR - Paris
- Henri Prade, *Directeur de Recherche CNRS*, IRIT - Toulouse
- Marie-Christine Rousset, *Professeur*, LIG - Grenoble
- Marc Schoenauer, *Directeur de Recherche INRIA*, INRIA - Paris
- Thomas Schiex, *Directeur de Recherche INRA*, INRA - Toulouse
- Cordelia Schmid, *Directeur de Recherche INRIA*, INRIA - Grenoble
- Jean-Philippe Vert, *Directeur de Recherche Mines ParisTech*, Mines ParisTech - ENS Paris
- Sabine Bruaux, *DGE*
- Amaury Fleges, *CGI*
- Claire Giry, *CGI*
- Benoit Le Blanc, *DGRI* (*Maître de Conférences*, IMS - Bordeaux)

Annexe 2 - Liste des contributeurs extérieurs au GT

Voici la liste des personnes, extérieures au groupe de travail recherche amont, ayant apporté des contributions à nos travaux:

- Gilles BAILLY, *ISIR*, Paris.
- Sébastien DESTERCKE, *HEUDIASYC*, Compiègne.
- Christine GASPIN, *INRA*, Toulouse.
- Yvonne JANSEN, *ISIR*, Paris.
- Nicolas SABOURET, *LIMSI*, Paris.
- Michèle SEBAG, *LRI*, Paris.

Annexe 3 – Carte de l'Intelligence Artificielle

Intelligence Artificielle		
acronyme	**thématique**	**conférences principales**
RCR	Représentation des connaissances et raisonnement	KR - IJCAI
INC	Incertitude - Décision - Réseaux Bayésiens - Modèles Graphiques	UAI - IJCAI
PPC	Contraintes - SAT	SAT - CP - IJCAI
SMA	Agents autonomes et Systèmes Multi-Agents	AAMAS - IJCAI
PRH	Planification - Recherche Heuristique	ICAPS - IJCAI
APP	Apprentissage automatique	ICML - NIPS - IJCAI
WEB	IA & Web	WWW - DL - IJCAI
TAL	Traitement Automatique des Langues	ACL
ROB	Robotique	
VISION	Vision - Reconnaissance des Formes	CVPR - ICCV
NEU	Modélisation cognitive et systèmes neuro-informationnels	NIPS

Autres Thématiques Informatiques Proches		
acronyme	**thématique**	**thématiques en lien avec l'IA**
IT	Informatique Théorique	Logique / Complexité Algorithmique / Théorie des graphes
BD	Bases de Données	Web sémantique / Big Data
RO	Recherche Opérationnelle	SAT / Contraintes / Décision / Recherche Heuristique
IHM	Interaction Homme-Machine	

Autres Disciplines en lien avec l'IA		
acronyme	**thématique**	**exemple d'interactions**
	Philosophie	Raisonnement
	Mathématiques (statistiques)	Apprentissage
	Neurosciences	Apprentissage (réseaux de neurones)
	Economie	Systèmes Multi-Agents
	Linguistique	Traitement automatique des langues
	Sciences cognitives	Systèmes Multi-Agents

Annexe 4 - Description des sujets de recherche disruptifs

Le groupe de travail recherche amont a identifié **36 sujets de recherche disruptifs.** Ces sujets ont été classés en **9 grandes thématiques**: **perception; collaboration humain machine; intelligence artificielle à large échelle; explicabilité; apprentissage machine; résolution de problèmes; décision, agents autonomes et intelligence collective; intelligence artificielle générale; enjeux éthique et sociaux, acceptabilité, et protection de la vie privée.**

Chacun de ces sujets de recherche disruptifs et de ces grandes thématiques est détaillé ci-dessous.

Perception

La perception et les informations perceptuelles sont une entrée de beaucoup de systèmes de raisonnement. Par exemple, pour concevoir des robots capables d'interagir avec leur environnement les capacités de perception sont le premier chainon indispensable avant de mener un raisonnement, de prendre une décision et d'agir. Or beaucoup reste à faire sur ces problématiques liées à la perception.

Les grands défis concernant cette thématique:

Vision

La recherche en vision artificielle connaît des développements sans précédent depuis une quinzaine d'années sur des aspects en lien avec l'intelligence artificielle. Cette réussite repose à la fois sur des avancées en termes de modélisation de l'information visuelle et de représentation des connaissances, sur des méthodes d'apprentissage statistiques exploitant des données massives annotées, et enfin sur des calculateurs très puissants adaptés aux types de calcul nécessaires (GPU). Les approches de modélisation reposant sur des architectures neuronales profondes, en particulier les réseaux de neurones convolutionnels, sont aujourd'hui les approches dominantes. De très nombreux champs s'ouvrent dans la discipline pour revisiter différentes tâches de vision complexes (comme la segmentation sémantique, la détection d'objets, la reconnaissance d'actions) en faisant collaborer ces modèles avec d'autres approches (comme les modèles graphiques ou plus généralement des méthodes par raisonnement). Il reste également des défis importants concernant la représentation des données visuelles (géométriques, structurées, etc.) ou multimodales (vision et langage), et l'apprentissage dans des contextes variées (faiblement ou non supervisé) afin de vraiment gagner en généralité dans les applications de vision. La génération automatique de données d'apprentissage est une question ouverte. Enfin, les méthodes associant apprentissage à partir des données, représentation des connaissances et raisonnement deviennent un défi important, par exemple pour la compréhension globale et de haut niveau de scènes visuelles (statiques, dynamiques, 3D) intégrant les relations et interactions entre objets.

Parole / Audio / Dialogue

Le traitement automatique du langage parlé (TALP) et le dialogue doivent résoudre des problèmes spécifiques liés à la nature du signal parlé ou de la communication interactive. L'approche dominante pour retranscrire le signal s'appuie sur un apprentissage statistique, en particulier les réseaux profonds, exploitant des données annotées. L'étape de transcription automatique précède généralement l'application de programmes de traitement automatique des langues dédiés à la compréhension, traduction, analyse de ces énoncés. Les principaux enjeux, outre l'amélioration des méthodes d'apprentissage aux

coeur des programmes de transcription automatique, portent sur la définition de méthodes d'analyses robustes qui soient capables, pour le langage parlé, de prendre en compte les erreurs générées par les reconnaisseurs de la parole, et pour le dialogue écrit ou oral ou écrit/oral, la caractère non standard de la langue (tweet, chat, forum), les niveaux de langue et les variétés dialectales; de méthodes multimodales n'effectuant pas de traitement en cascade mais des analyses conjointes de plusieurs modalités; de nouveaux modèles permettant de représenter et exploiter les caractéristiques des énoncés linguistiques et du signal oral, et plus largement le signal de l'écriture manuscrite, l'image, la vidéo; de modèles et méthodes pour la prise en compte du contexte de production du message oral ou écrit de manière à adapter l'analyse automatique et mieux prendre en compte l'utilisateur final.

Multimodalité

La communication est intrinsèquement multimodale. La reconnaissance des intentions communicatives, dans les sens humain-machine ou machine-humain, inclut la parole, les signaux acoustiques et visuels, les expressions, les émotions et autres signes sociaux comme les attitudes ou signes d'engagement. Des modèles d'analyse et de reconnaissance spécifiques à chaque modalité doivent être développés. L'interprétation des signaux en termes de fonctions socio-émotionnelles requiert leur fusion, mais aussi la prise en compte du contexte.

Collaboration Humain Machine

Cette problématique est primordiale pour améliorer les interactions humains-machines, afin de permettre aux utilisateurs d'avoir une meilleure interaction avec les programmes et robots. Cela passe en particulier par la facilitation de l'interaction et donc le développement d'outils pour le traitement du langage naturel.

Les grands défis concernant cette thématique:

Recommandations à l'utilisateur

Les systèmes de recommandation cherchent à prédire l'évaluation qu'un client potentiel peut faire d'un produit, sur la base des évaluations données par le client et d'autres clients sur différents produits (le système recommande ensuite les produits dont l'évaluation prédite est élevée). Un problème auquel se heurtent ces systèmes est celui de la "sérendipité" (c'est-à-dire la capacité à proposer des choses nouvelles, plutôt que seulement des produits trop semblables à ceux déjà bien notés par le client). Ces nouveaux produits que le système serait capable de proposer à bon escient pourraient gagner aussi à être expliqués en référence aux choix précédents du client, et à ce qu'on a pu apprendre de son profil de préférences.

Décision sous incertitude

Certaines décisions sont non répétées et ne doivent donc pas seulement être bonnes en moyenne. De plus, l'incertitude peut présenter un aspect épistémique (dû à une information incomplète), sans qu'il s'agisse de se protéger contre la variabilité d'un phénomène. La décision face à l'incertitude doit alors prendre en compte l'attitude pessimiste, ou optimiste du décideur. L'incertitude peut aussi n'être appréciée que de manière qualitative. Un système d'aide à la décision devrait si nécessaire, être capable d'une manière générale, d'expliquer, d'argumenter les décisions qu'il propose. Il devrait aussi pouvoir

proposer des options susceptibles d'être intéressantes pourvu que certaines actions complémentaires accompagnent la décision. Ceci vaut pour des décisions isolées, mais aussi pour des plans d'actions.

Sémantique et compréhension de la langue

Le processus de transfert de la syntaxe vers la sémantique et inversement de la sémantique vers syntaxe nécessaire à la compréhension et la génération d'un message nécessite l'adoption de modèles de représentation de la sémantique des langues. Ces modèles relèvent soit de la sémantique formelle qui rencontrent des problèmes de robustesse, soit de la sémantique distributionnelle qui souffrent d'inexactitude. Les enjeux portent sur la proposition de modèles formels robustes et distributionnels exacts, mais aussi sur la convergence des sémantiques compositionnelles et distributionnelles en proposant des modèles intégratifs; de méthodes pour transformer des données langagières modélisées comme les ontologies en texte et inversement; de modèles et de méthodes qui permettent l'interprétation du texte, et en particulier l'analyse ou la génération d'un discours.

Intentionnalité / Emotions

La prise en compte des émotions et des intentions des humains dans une interaction demande leur détection et reconnaissance chez l'humain à partir de signaux multimodaux (e.g., acoustiques, visuels, parole), leur simulation par la machine (comme par le comportement non-verbal d'un robot, la voix de synthèse émotionnelle) et l'adaptation de la machine aux émotions et intentions des utilisateurs. Les premières tâches demandent des méthodes d'analyse de signaux multimodaux et de modèles de fusion (pour la reconnaissance) et de fission (pour la simulation). Une difficulté majeure réside en la très large variété d'émotions et d'intentions dans une interaction. Les modèles d'adaptation reposent généralement sur des modèles de raisonnement du processus cognitif de l'humain et des modèles de stratégies de dialogue.

Intelligence Artificielle à large échelle

Une des données saillante à l'heure actuelle est le fait que dans de nombreux domaines une grande quantité de données est disponible, et il faut adapter les méthodes usuelles (raisonnement, décision, apprentissage, etc) pour pouvoir appréhender efficacement ces données.

Les grands défis concernant cette thématique:

Raisonner sur des données massives

Les algorithmes de raisonnement mécanisent des inférences à partir d'ensembles d'expressions symboliques, dans le cadre de formalisations logiques. La complexité du raisonnement dépend de l'expressivité du langage dans lequel les formules sont exprimées, combinée avec la taille des données correspondant à des formules très simples mais pouvant être en grand nombre. Il est important de développer de nouvelles formes de raisonnement donnant une plus large place aux données et passant à l'échelle de données de grande taille. Pour cela, il est essentiel de revisiter les différentes formes de raisonnement et les formalismes de représentation de connaissances. Le développement d'approches de raisonnement décentralisé sur des données distribuées est en particulier un défi important à relever.

Intégration de données hétérogènes et/ou multisources

Dans beaucoup de domaines d'application, de nombreuses données sont récoltées et disponibles dans différentes sources de données ayant une organisation et une structure qui leur sont propres. L'interrogation unifiée de sources de données multiples et le liage de données hétérogènes sont des défis majeurs pour permettre un accès uniforme à des usagers (décideurs, experts, grand public) ayant des besoins d'analyse très variés. Il s'agit de développer d'une part les méthodes de synthétisation (fusion) de ces données, mais également des approches à base de connaissances qui permettent de garder la provenance des données dans les résultats d'analyse obtenus de sorte que les sources et liens puissent être vérifiés et interrogés et ainsi rendre les résultats plus compréhensibles et donc mieux acceptés par un utilisateur non informaticien.

Aspects matériels et calcul (architectures distribuées, etc.)

Le traitement de grandes quantités de données créé des contraintes matérielles qui ne peuvent être ignorées par les algorithmes mis en oeuvre. Par exemple, un stockage décentralisé des données sur de multiples serveurs nécessitera des algorithmes décentralisés qui puissent traiter localement des sous-ensemble des données, et communiquer un minimum d'information entre les serveurs. La prise en compte de telles contraintes pose de nouveaux défis mathématique et algorithmiques.

Explicabilité

Une capacité importante pour les systèmes artificiels automatisés est le fait de pouvoir expliquer, argumenter, justifier un raisonnement, une décision ou une action. Cela est indispensable pour une meilleure acceptabilité des programmes prodiguant conseils ou injonctions (dans le cas d'un véhicule autonome par exemple).

<u>Les grands défis concernant cette thématique:</u>

Résumé / Restitution linguistique

La restitution linguistique porte sur la production d'un contenu linguistique adoptant un discours précis, répondant à un besoin d'intervention, de diagnostic ou de synthèse, et généré pour un utilisateur. Les formes de restitution sont multiples comme les réponses rédigées à une question précise, les résumés d'un article, d'un livre ou d'un ensemble de tweets, les recommandations, les tendances, les opinions, ou encore les exercices de validation de connaissances. La restitution linguistique s'appuie sur des techniques d'extraction d'informations employant principalement des méthodes d'apprentissage supervisé. La restitution linguistique devrait s'enrichir en se couplant à d'autres types de restitution (graphique, image), se dynamiser pour intégrer l'évaluation perpétuelle des contenus au cours du temps et s'incrémenter des connaissances validées par les humains.

Argumentation et Délibération

L'argumentation est un processus standard pour trouver une position commune dans un groupe à partir d'opinions individuelles contradictoires. Les plate-formes permettant d'avancer des arguments, des attaques ou des soutiens entre arguments se développent de plus en plus sur le Web. L'argumentation est également de plus en plus souvent utilisé pour les consultations citoyennes (voir l'exemple de la loi numérique). Il est nécessaire d'avoir les méthodes de raisonnement et de décision pour pouvoir exploiter ces arguments et permettre par exemple d'identifier les arguments les plus acceptables, ceux qui sont le plus polémique, etc. Ces méthodes sont étudiées depuis de nombreuses années et les bases sont bien connues, mais ces nouveaux usages et leur conséquences (grandes masses d'arguments, nouveaux types de relations entre argument, etc.) nécessitent de trouver de nouvelles méthodes afin de pouvoir faciliter le développement et l'utilisation de ces plate-formes de délibération, qui auront un rôle clé à l'avenir pour les délibérations et la consultation citoyenne, et les applications de commerce (avis sur les produits, etc.).

Explicabilité des méthodes d'apprentissage

Le fait qu'apprentissage et capacité d'explication ne vont en général pas de pair tient, au delà d'une antinomie entre performance et explicabilité, à l'histoire même de l'IA où dès ses débuts plusieurs courants de recherche bien distincts se sont développés : d'une part des méthodes numériques de type "boite noire", et d'autre part des méthodes basées sur la logique orientées vers le raisonnement. Cela vaut aussi pour l'apprentissage à partir de données booléennes, nominales, ou ordinales. Il existe cependant un petit nombre de travaux qui interfacent de telles "boites noires" avec des systèmes de règles qui peuvent éclairer leurs réponses. Développer de telles interfaces explicatives est indéniablement un enjeu de grande importance pour le futur.

Causalité

La causalité a déjà fait l'objet de nombreuses études que ce soit pour le raisonnement symbolique ou probabiliste, mais il n'y a pour le moment pas de solution parfaite. Or ces problématiques de causalité et de responsabilité sont de première importance pour de nombreuses questions légales (par exemple identifier les responsabilités lors d'un accident de la route), ou pour prédire les conséquences d'une action comme changer un paramètre d'un système. Des travaux récents en apprentissage ont proposé de nouvelles approches pour identifier la causalité quand on observe des corrélations, mais restent assez peu performants en pratique; la causalité pose des questions mathématiques et algorithmiques difficile mais de première importance.

Apprentissage machine

La puissance des techniques d'apprentissage est illustrée par le déploiement de celles-ci pour un nombre impressionnant d'applications (notamment la vision, la traduction automatique, les véhicules autonomes, etc.). Ces techniques sont applicables principalement lorsque l'on dispose de masses de données. Mais il reste à travailler sur un grand nombre de points, que ce soit sur les fondements de ces approches, ou sur les cas où l'on dispose de moins d'exemples.

<u>Les grands défis concernant cette thématique:</u>

Apprentissage profond

Au sein de la révolution que connait l'Intelligence Artificielle aujourd'hui, l'apprentissage profond (deep learning) occupe une place de tout premier plan. Il est au coeur des technologies qui permettent aujourd'hui de réaliser des tâches encore impensables il y a quelques années, comme la conduite automatique ou la conversation avec des assistants vocaux. La réussite du deep learning repose à la fois sur des avancées en terme de modélisation des représentations de l'information, sur des méthodes d'apprentissage statistiques exploitant des données massivement annotées, et sur des calculateurs très puissants adaptés à ces architectures.
Parmi les enjeux en recherche essentiels sur le deep learning, on trouve l'enjeu d'une meilleure compréhension théorique du fonctionnement de ces réseaux profonds ; des avancées sur les méthodes d'optimisation pour faciliter l'apprentissage de ces réseaux ; la simplification et/ou compression de ces réseaux ; l'adaptation de ces architectures pour des modalités d'apprentissage faiblement ou non supervisées, du multi-taches.
Ces enjeux sont de taille tant les applications industrielles et les implications sociétales sont immenses aujourd'hui.

Apprentissage à partir de peu d'exemples / non supervisé

La plupart des applications réussies récentes de l'apprentissage supervisé (comme en traitement automatique des langues et en vision artificielle) ont été rendues possibles grâce à un grand nombre d'exemples d'entrée/sortie observées. Cependant, de nombreux nouveaux problèmes ont soit peu d'exemples pour entraîner les modèles, soit beaucoup de données mais la plupart partiellement (voire nullement) etiquetées (par exemple, des images sans descriptifs de leurs contenus). Ces deux déviations du cadre dominant restent des sujets très largement ouverts.

Apprentissage en ligne et incrémental

L'apprentissage est dit "en ligne" ou "incrémental" lorsque les données utilisés pour entrainer un modèle arrivent séquentiellement. Le modèle doit alors être mis à jour itérativement, au fur et à mesure de l'arrivée des données. Ce scénario a vocation à jouer un rôle central à l'avenir, avec le développement de capteurs connectés ou la récolte de données sur le web qui produisent des flux continus de données. La théorie et les algorithmes de l'apprentissage en ligne restent peu développées, notamment dans le cas où le flux de données n'est pas stationnaire et ou le modèle doit évoluer au cours du temps.

Apprentissage par renforcement

Alan Turing avait déjà, avec sa P-machine (que l'on peut éduquer à l'aide de 2 signaux, Pain and Pleasure) proposé la vision à long terme de l'apprentissage par renforcement (RL): apprendre à choisir l'action qui donnera à terme la plus grande récompense.
Déjà performants il y a 20 ans (champion de backgammon, 92), les algorithmes de RL ont vu leurs performances décuplées par l'arrivée des algorithmes dits de bandits manchots (incontournables dans le domaine de placement de publicité, avec l'impact que l'on sait sur le business model du Web), et de

l'apprentissage profond, permettant d'utiliser tout simulateur un tant soit peu robuste pour générer des masses de données d'apprentissage. Les réussites spectaculaires dans le domaine des jeux vidéos (à partir des images de l'écran), puis du GO (couplage bandits manchots / apprentissage profond) ne doivent pas masquer d'autres domaines où l'impact du RL est en plein essor, comme les politiques d'énergie, ou la robotique, où l'on retrouve le rêve de robots compagnons que l'on pourrait éduquer comme des animaux (sans spécifier précisément la récompense). A l'opposé, l'autre enjeu majeur reste l'obtention de garanties théoriques.

Transfert et Analogie

L'apprentissage par transfert consiste à exploiter un modèle appris sur une tâche donnée pour une autre tâche, si possible similaire. Ce scénario prend toute son importance lorsque l'on désire résoudre un grand nombre de problèmes mais que l'on ne dispose pas de beaucoup d'exemples sur tous les problèmes, par exemple, pour développer des systèmes de reconnaissance vocale dans de nombreuses langues différentes. Ce transfert est de nature analogique. De manière plus générale les rapports analogiques pourraient être davantage exploités en apprentissage, en particulier pour apprendre à partir de peu d'exemples.

Résolution de Problèmes

Le machine learning permet, mieux que jamais, d'aborder la résolution de problèmes qui nous semblent intuitifs (reconnaître des entités sur des images) au travers de grandes quantités d'exemple annotés. Mais une IA doit aussi être capable de résoudre des problèmes de raisonnement logique bien posés difficiles, pour l'être humain comme pour la machine, qu'il s'agisse de prouver des théorèmes ou de résoudre des casse-têtes de tout ordre, y compris ceux d'intérêt pratique…

<u>Les grands défis concernant cette thématique:</u>

Problèmes NP-difficiles (Beyond NP)

La résolution pratique d'instances de problèmes de raisonnement logique dans des espaces discrets, bien que (NP)difficile en théorie, a fait des progrès considérables sur les 20 dernières années. Qu'il s'agisse de formules logiques propositionnelles, de réseaux de contraintes ou d'autres modèles graphiques, les algorithmes et outils de résolution (eg. de ``clause learning''), sont maintenant capables de résoudre exactement des problèmes industriels de plus d'un million de variables booléennes. Cette capacité sur-humaine, encore mal comprise théoriquement, a déjà révolutionné l'industrie de la vérification des circuits intégrés (processeurs, Intel en particulier). Plus proche de nous, elle est aussi utilisée pour concevoir les problèmes de Sudoku du journal ``Le Monde'' et forme aussi le coeur du programme français ``Woodstock'' qui a gagné la première place lors de la compétition de ``General Game Playing'' (Stanford, 2016). La France est un leader international effectif du domaine, aussi bien en termes de publications (cf. http://www.a4cp.org/cparchive/countries_by_year) que d'outils de résolution.

Les grandes questions qui se posent à présent sont de comprendre la raison de l'efficacité des méthodes développées dans les 20 dernières années; de les combiner avec du machine learning pour résoudre des problèmes difficiles mais définis seulement par des données ; de les utiliser pour résoudre des problèmes difficiles d'intérêt pour le machine learning et au delà (calcul de marginales, comptage, programme "Beyond NP" proposé aux USA par le président de l'ACM) ; d'utiliser le machine learning pour rendre ces algorithmes adaptatifs ; et de les combiner avec de l'optimisation continue.

Optimisation

Un bon nombre de sous-problèmes fondamentaux en machine learning (optimisation/estimation de paramètres avec une fonction de perte pénalisant la complexité) tout comme en raisonnement logique difficile (calcul de bornes et inférences massives efficaces) se réduisent in fine à des problèmes d'optimisation continue ou de satisfaction discrète en grande dimension, avec des formes de critères spécifiques, éventuellement bruités. Tout progrès marqué sur l'efficacité théorique ou pratique des algorithmes d'optimisation ou l'identification de nouveaux cas traitables efficacement (classes polynomiales) significatifs peut avoir un effet massif sur les capacités des outils de l'IA, en machine learning comme en raisonnement logique.

Certaines architectures de calcul (quantiques, photoniques) dont on voit des prototypes industriels au niveau international (eg. D-Wave), pourraient potentiellement changer la donne car elles sont ciblées sur la résolution efficace de problèmes d'optimisation et de comptage discret. Il s'agit d'une thématique portée par la physique quantique.

Recherche heuristique et méta-heuristique

Dans le domaine de l'optimisation (et plus particulièrement l'optimisation dans des espaces discrets, mixtes, ou peu structurés), et par opposition à la recherche exacte, la recherche heuristique applique des règles d'exploration de l'espace qui sont spécifiques au type de problème (appelées heuristiques). Les méta-heuristiques sont des façons génériques d'appliquer ces règles spécifiques, et de sélectionner les solutions que l'on va conserver d'une itération à l'autre - souvent inspirées librement par des phénomènes naturels (recuit simulé, évolution et sélection naturelle, colonies de fourmis, etc). La souplesse d'adaptation de ces algorithmes à des problèmes ne tombant pas dans les formalismes standards des problèmes combinatoires leur a permis d'obtenir de nombreux succès industriels ... et quelques succès académiques. Les sujets de recherche disruptifs sont la recherche de résultats théoriques à la hauteur des réalisations pratiques, au moins pour les problèmes combinatoires classiques, la construction automatique de méta-heuristiques ad hoc (oxymore) pour des classes de problèmes, voire par instance, à partir de descripteurs des instances, et l'hybridation de méthodes exactes et heuristiques

Raisonnement Multi-Echelle

Le monde dans lequel évoluent des agents humains ou artificiels peuvent être décrits, appréhender à différents niveaux de granularité, selon le niveau d'abstraction nécessaire pour l'évaluation, l'action ou l'explication. Il n'existe pour ainsi dire pas de systèmes capables de passer d'un niveau de description à un autre selon les besoins. Or, cela semble nécessaire, à la fois en terme d'explicabilité (pour intéragir avec des utilisateurs), et en termes de la réduction potentielle de la complexité des problèmes.

Décision, agents autonomes et intelligence collective

Il y a beaucoup d'applications potentielles demandant l'intervention de plusieurs agents, qui peuvent être des entités artificielles logicielles (sur internet) ou matérielles (robots). On peut imaginer par exemple le développement d'assistants personnels représentant et négociant pour leur propriétaire, ou des applications où plusieurs robots doivent collaborer pour le secours de personnes, etc. Il est crucial de disposer des outils formels et algorithmiques nécessaires à la modélisation et à la régulation de ces

interactions. Certaines des recherches dans ce domaine font appel à des notions et techniques venant de la microéconomie (en particulier théorie des jeux et théorie du choix social).

Choix social computationnel / Théorie des jeux

Il faut travailler au rapprochement, suivi si possible, d'une véritable collaboration, avec des disciplines connexes : économie, systèmes complexes, sociologie, biologie des populations, robotique. Dans le cas particulier du choix social computationnel (pour la décision collective), un enjeu important est de convaincre les collectivités publiques d'utiliser les outils développés par la communauté, et d'utiliser leur feedback pour commencer d'autres développements. Ce feedback sociétal est un véritable sujet disruptif. Il a commencé à être pris en compte, notamment par la plateforme de partage équitable en ligne Spliddit (http://www.spliddit.org/).
Les applications de la thématique "économie computationnelle" en commerce électronique sont clairs: systèmes de recommandation, systèmes de réputation (par exemple, TripAdvisor), systèmes d'échange (P2P ou pas) et problèmes de confiance, systèmes d'enchères sur le web. A noter que la France n'est pas très active sur ces thématiques (à part les systèmes de recommandation).

Raisonnement sur l'autre

Un agent autonome doit souvent être capable de raisonner sur les croyances, préférences et intentions d'autres agents (que ce soient des humains ou d'autres agents artificiels), soit parce qu'il est nécessaire d'agir contre d'autres agents, ou en collaboration avec d'autres agents, ou, enfin, pour d'autres agents. Les grands défis sont ici la décision en contexte hostile ou compétitif : jeux de sécurité ("security games"), ou dans un contexte plus anodin, programmation de joueurs (échecs, go, poker, et aussi jeux vidéos). Dans tous ces contextes il est crucial d'avoir un modèle pertinent des croyances, préférences et intentions de l'opposant. La recherche doit se faire de façon de plus en plus pluridisciplinaire : machine learning, recherche heuristique, théorie des jeux, économie expérimentale, psychologie cognitive.
Le problème se pose également pour la décision en environnement collaboratif, où une équipe d'agents (par exemple, des robots) doit résoudre des tâches collectivement. Il s'agit de résoudre de manière décentralisée un problème qui se pose globalement à la collectivité d'agents (et donc aux humains qui les ont programmés). Un exemple typique d'application : la résolution d'une tâche d'exploration dans un environnement mal connu par une flotte de robots se déplaçant de façon autonome et pouvant (plus ou moins bien) communiquer entre eux. On est ici en plein dans le cadre de l'IA distribuée, avec des intersections fortes avec les domaines de la planification, du raisonnement et de la décision dans l'incertain, voire de la robotique.
Enfin, une dernière application importante, est l'assistance à la personne, où il s'agit cette fois de modéliser les états mentaux de la personne assistée (personne âgée ou en situation de handicap). Là encore, les sujets en rupture sont hautement pluridisplinaires.

Actions conjointes / Planification collaborative

La coopération/collaboration entre agents, humain-machine (en particulier humain-robot), est à la fois physique et cognitive. Pour réaliser des tâches et actions conjointes, les agents doivent partager des informations sur l'environnement, sur leurs objectifs communs ou individuels, et sur la manière d'accomplir les tâches. Ils doivent aussi maintenir une communication leur permettant d'inférer que le co-acteur est bien engagé dans la tâche commune. Le robot doit donc élaborer et maintenir un modèle

de l'humain, de ses connaissances et de ses intentions, ce qui exige une capacité de prise de perspective et de raisonnement sur autrui. La planification, avec des contraintes temporelles, des actions de la machine doit inclure aussi celles qui seront effectuées par l'autre agent ainsi que les actions de coordination.
L'obstacle majeur pour la plupart des sujets de recherche concernés est le passage à l'échelle. C'est très nettement le cas pour la résolution collective de problèmes: la planification d'un seul agent dans un environnement partiellement connu est déjà un problème très difficile, et l'introduction de l'aspect distribué rend les choses encore bien plus complexes, notamment en raison de la prise en compte des croyances des agents sur les plans et croyances des autres agents, et de la nécessité de planifier des actions de communication visant à réduire l'incertitude.

IA intégrative et robotique

Le système nerveux et le cerveau ont évolué vers ce qu'ils sont chez l'être humain aujourd'hui en développant des capacités de perception, de raisonnement, d'apprentissage et de communication, pour une action plus efficace et plus adaptée. Le robot comme paradigme de l'IA matérialisée, agent à la fois autonome et interactif, est soumis à la complexité et à l'incertitude du monde réel dans lequel il évolue et dont il doit respecter sa dynamique. Il doit intégrer, en tant que système, des capacités de perception, d'action, de raisonnement, d'apprentissage, et de communication/interaction dans une architecture de contrôle globale ou architecture cognitive. Il est important d'observer que l'étude de ces capacités de manière séparée risque d'aboutir à des solutions inappropriées à leur intégration, même si elles sont individuellement performantes. En revanche la prise en compte conjointe de leurs liens aboutit à des approches qui doivent nécessairement appréhender le réel de manière opérationnelle, et par conséquence la problématique de recherche elle-même s'en trouve renouvelée.
L'apprentissage par renforcement est un apprentissage non supervisé qui permet au robot de découvrir à la fois les effets de ses actions et l'incertitude caractérisant ses actions ainsi que l'environnement. Le lien entre perception, action et apprentissage - en particulier apprentissage par renforcement - permet d'associer les représentations perceptuelles aux capacités d'action pour extraire une sémantique du monde en particulier à travers la notion d'affordance qui rend compte des propriétés des objets en ce qu'elles représentent pour l'agent. D'autre part cette interaction pose la question de l'apprentissage à partir d'une faible quantité de données ou d'expériences.
La prise de décision dans un système intégré nécessite l'étude de l'architecture qui permet la mise en œuvre des différentes fonctionnalités de manière cohérente et coordonnée, au delà des problématiques du middleware. Il faut comprendre comment intégrer des processus décisionnels de complexités différentes permettant la délibération et la réactivité à travers des mécanismes réflexifs comme le méta-raisonnement et l'auto-évaluation.

Intelligence artificielle générale

Au delà du développement nécessaire de techniques dédiées à des problèmes particuliers, il est important de développer des techniques permettant de s'adapter à des situations peu ou pas connues à l'avance. Ce que nous avons résumé en ``intelligence artificielle générale".

<u>Les grands défis concernant cette thématique:</u>

Informatique neuro-inspirée

L'informatique classique s'est développée sur le modèle de la machine de Turing dans laquelle mémoire et instructions sont physiquement séparées. Une autre informatique, dans son sens le plus géné-

ral de science de l'information, peut être imaginée en se référant aux principes essentiels de la cognition mentale, lesquels restent à découvrir. La façon dont l'information est matérialisée, remémorée, propagée et exploitée dans les circuits corticaux pourra alors servir de nouveau paradigme dans l'élaboration d'une intelligence artificielle généraliste.

IA embarquée

Les réalisations récentes les plus marquantes de l'intelligence artificielle, telles que les assistants personnels (Siri, OK Google, Cortana, …) ou les machines à jouer (AlphaGo, DeepStack, Watson, …) utilisent une puissance de calcul qui ne peut, en l'état actuel de la technologie, être embarquée dans des dispositifs autonomes. Beaucoup d'applications de l'IA, civiles et militaires, seront cependant appelées à fonctionner sous contraintes fortes d'autonomie (absence de communication), de temps de réponse et de confidentialité. De nouvelles architectures de circuits, voire de nouveaux procédés, adaptés aux systèmes embarqués, restent à imaginer.

Créativité et génération

Le problématique de la créativité, longtemps cantonnée à des études psychologiques est devenue une problématique importante, du fait du progrès considérable des méthodes d'apprentissage machine, et des applications nombreuses dans l'industrie du loisir (génération de jeux vidéos, de scripts, de musique, texte, etc.). Il s'agit de modéliser des systèmes capables de produire/générer/engendrer des artefacts de divers natures (images, son, musique, textes, etc.) capables de rivaliser avec les productions humaines correspondantes.

Jeux génériques (General Game Playing)

La problématique des jeux génériques est de concevoir des méthodes permettant de jouer à des jeux non déterminés à l'avance. Contrairement aux méthodes dédiées pour les échecs, le go, le poker, etc., qui sont optimisées dans un but unique, la question ici est d'avoir une méthode générique, permettant de jouer rapidement (il n'y a pas de grands jeux de données et la méthode n'a que quelques minutes pour analyser les règles qu'on lui envoie pour décrire le jeu avant de commencer à jouer). Cette problématique nécessite l'intégration de beaucoup de méthodes d'intelligence artificielle différentes (raisonnement, apprentissage, contraintes, etc.), et il s'agit, dans un environnement suffisamment contrôlé, de développer des méthodes permettant de s'adapter à des problèmes non spécifiés à l'avance, ce qui est une capacité importante à maitriser si l'on veut avoir dans le futur des agents/robots pouvant s'adapter à des problèmes et à des environnements peu connus (et sur lesquels il n'est pas possible de s'entraîner à l'avance).

Intégration de différents modes de raisonnement

Après les premiers succès des systèmes experts il y a trente ans, et la prise de conscience ensuite de leurs limitations, nombre de travaux en IA ont permis des avancées considérables sur une grande variété de problèmes de raisonnement: raisonnement dans l'incertain, raisonnement en présence d'exceptions, raisonnement en présence d'incohérences et fusion d'informations provenant de plusieurs sources, raisonnement à partir de cas, diagnostic, raisonnement interpolatif, raisonnement sur le temps, sur l'espace, raisonnement déontique, raisonnement sur l'action, raisonnement sur les préfé-

rences, planification, raisonnement épistémique multi-agents, raisonnement argumentatif, raisonnement sur les émotions, etc. Il n'existe cependant pour ainsi dire pas de système capable d'utiliser de manière intégrée plusieurs de ces modes de raisonnement. Ceci s'avérerait pourtant très utile si on veut développer des systèmes d'aide à la personne, d'accompagnement, d'assistance de diverses sortes, où il est nécessaire de combiner les capacités de plusieurs de ces types de raisonnement.

Enjeux éthique et sociaux, acceptabilité, et protection de la vie privée

Il sera nécessaire de travailler sur les enjeux éthiques et sociaux des méthodes d'intelligence artificielle. Si une grande partie de ces considérations concernent directement l'éthique et les sciences sociales, et sortent donc du cadre de ce groupe sur la recherche amont en IA, certaines de ces questions sont des sujets de recherche à part entière en IA.

Les grands défis concernant cette thématique:

Intégration de la préservation de la vie privée dans les méthodes d'apprentissage et de raisonnement

Il s'agit ici de promouvoir des recherches sur les modèles mathématiques et informatiques de préservation de la vie privée, et leur intégration ``par construction'' dans les algorithmes d'apprentissage automatique, de raisonnement à partir de données, et plus généralement dans les méthodes d'intelligence artificielle se fondant sur les données.

L'explosion des données personnelles sur la toile et des traces enregistrées par les ordinateurs sur les activités des personnes, combinée à des algorithmes d'IA de plus en plus puissants, ouvre des perspectives sans précédent sur la détection de risques, la médecine préventive, et l'aide à la personne; mais augmente considérablement le risque de violation de la vie privée. Des résultats récents ont montré que le croisement entre jeux de données même anonymisées peut permettre de ré-identifier des personnes. Derrière cette question, se cachent en réalité différents verrous scientifiques : dont (i) la définition de modèles de préservation de la vie privée selon le but et la nature des données (distribuées, structurées, grandes dimensions, etc), (ii) la définition de formalismes et boites à outils bien fondés mathématiquement pour l'étude des garanties théoriques de ces modèles (compromis confidentialité/précision par exemple).
Les communautés d'apprentissage automatique et de bases de données se sont emparés de ces verrous ces dernières années, au travers notamment du modèle de confidentialité différentielle.

Ce sont là des efforts à encourager, en créant les conditions d'une fertilisation croisée entre travaux en cryptologie, apprentissage et raisonnement à partir des données, algorithmique et complexité, et statistique et probabilités. L'objectif visé doit être que tout algorithme d'IA fondée sur des données vienne avec des garanties fortes de confidentialité (à l'instar de garanties de convergence pour les algorithmes d'apprentissage statistique). Il est là question de confiance dans les usages de ces approches, qui si elle est renforcée favoriserait la mise à disposition de données sensibles (génomes par exemple) dont l'exploitation est porteuse de grandes promesses en termes de découvertes scientifiques et de retombées sociétales.

Acceptabilité et appropriation

A chaque rupture technologique doit être imaginée la rupture de mentalité qui ira avec. La capacité de pouvoir modifier une image avec nos doigts a instauré en nous un nouveau rapport à l'image, un lien direct entre vision et toucher. Avec l'évolution des recherches et des réalisations liées à l'intelligence artificielle, nos rapports avec la voix, avec la mémoire, avec la décision vont tout autant se retrouver bouleversées. Prendre seulement en considération le volet "acceptation" d'une technologie, revient à réduire le couplage (Usagers, Technologies) à la simple considération d'un comportement face à un dispositif. Pour dépasser cela, il ne faut pas faire l'impasse sur les intentions, les détournements, les envies. Plutôt que de parler d'acceptation des technologies, il faut envisager d'emblée l'appropriation de ces technologies. Ce glissement de l'acceptabilité vers l'appropriation demande l'éclairage des sciences du comportement et une collaboration effective entre chercheurs en IA et chercheurs en sciences humaines et sociales. La co-construction de solutions aboutit à des résultats plus performants ou tout simplement plus adaptés. C'est le cas dans les recherches sur la cobotique, sur la sociologie computationnelle, sur la cognitique ou encore sur les humanités numériques. Dans tous ces domaines, les études transdisciplinaires permettent d'anticiper ou de comprendre les évolutions de mentalités qui s'opèrent avec l'usage et l'appropriation des technologies issues de l'intelligence artificielle.

Aspects déontiques (respect des lois par les méthodes d'intelligence artificielle)

Les questionnements sur les problèmes éthiques, légaux et sociétaux (dits ELS) dans l'usage de l'IA et des robots datent d'une quinzaine d'années et sont devenus dernièrement assez prégnants avec l'apparition de nouvelles applications. Ils concernent des domaines très variés comme la protection de la vie privée, la surveillance, l'interaction avec des personnes vulnérables, la responsabilité morale et juridique, l'imitation du vivant, le statut du robot dans la société, le statut de l'humain augmenté par le robot, la prise de décision autonome (par exemple pour les voitures autonomes ou pour les armes), ou les risques hypothétiques posés par l'IA "générale". Les recherches sur ces différents aspects doivent impliquer des travaux interdisciplinaires associant par exemple les chercheurs en IA et robotique avec des philosophes, juristes ou psychologues. Mais posent également des questions purement théoriques, comme par exemple la validation/certification de méthodes développées en intelligence artificielle (i.e. peut-on prouver (vérifier) qu'une méthode satisfait un ensemble de lois données ?).

IA pour le droit (modélisation du raisonnement juridique)

Il existe de nombreux travaux théoriques concernant la représentation des connaissances et le raisonnement dans le contexte juridique, en utilisant différents types de logiques. Cependant, jusqu'à présent, seul un nombre limité de travaux consacrés aux systèmes juridiques ont été capables de simuler le raisonnement des juges leur permettant de prendre des décisions indépendamment de domaines d'application particuliers et circonscrits de la loi. Pour autant, la possibilité d'un mécanisme de raisonnement juridique générique (intégrant les différents types de raisonnement utilisés par les juges lors de la prise de leurs décisions) pour développer des systèmes juridiques (fondés sur l'IA) génériques est une perspective réelle. Ces systèmes pourraient alors être instanciés afin de traiter différentes situations d'application du droit. Cet axe de recherche comprend des dimensions liées à l'acceptabilité, à la déontique et à l'explicabilité mentionnées ci-avant. (voir rapport du sous-groupe dédié en annexe 6)

Annexe 5 - Description des grands domaines de recherche en IA

Nous donnons ici un panorama des grands domaines de recherche en IA. Le but étant de proposer une description synthétique et vulgarisée des différents domaines, de leurs buts, et de leurs enjeux actuels. Pour une description plus fine (et plus précise) de la discipline et de ces différents domaines on pourra se référer à *« Panorama de l'Intelligence Artificielle »*, édité par Pierre Marquis, Odile Papini et Henri Prade, chez Cepadues, qui comporte 3 volumes et 42 chapitres.

Apprentissage automatique (machine learning)

Les techniques d'apprentissage automatique sont aujourd'hui partie prenantes dans la plupart des succès médiatiques de l'IA, de la victoire de Watson à Jeopardy (même si la partie apprentissage ne constitue qu'une petite partie du système complet) aux succès retentissants de l'apprentissage supervisé en reconnaissance d'images et de vidéo ces dernières années, et de l'apprentissage par renforcement au jeu de Go et au Poker ces derniers mois. Si ces techniques existaient depuis quelques décennies déjà, c'est bien l'arrivée de l'apprentissage profond dans le paysage de l'IA récente qui a permis ces succès retentissants.

Apprentissage supervisé

Evolution scientifique depuis 2000 pour l'apprentissage supervisé
-Les années 2000 ont été dominées en large partie par des méthodes basées sur des problèmes d'optimisation bien posés (le plus souvent convexe), comme les méthodes à noyaux (de type SVM) ou plus généralement les techniques de régularisation par des normes. Les algorithmes viennent avec des garanties théoriques précises venant souvent des statistiques et fonctionnent sur des problèmes linéaires de grande taille (exemple : la publicité en ligne) ou non- linéaires de taille moyenne (exemple : vision avec peu d'images annotées).
-Les années 2010 ont vu l'émergence de l'apprentissage non-linéaire « end-to-end » à partir de bases de données beaucoup plus grandes et des architectures de calcul efficaces, l'exemple le plus marquant étant les réseaux de neurones profonds. Les garanties théoriques sont minces, mais la performance dépasse souvent l'état de l'art, un des atouts de ces méthodes étant leur scalabilité, en particulier grâce à des algorithmes sur des architectures distribuées (GPUs).

Evolution de l'environnement. Au niveau scientifique, la communauté voit naturellement vivre, s'éteindre, revivre des méthodes et des modes depuis 20 ans. De « big data », on est passé par « data science » puis maintenant « intelligence artificielle » pour parler d'un domaine dont les fondamentaux scientifiques ont peu évolué.
Un grand changement récent est l'intérêt porté par les entreprises (et notamment les GAFA), qui « grâce » à leur recrutement agressif de chercheurs et leur communication aussi très agressive, ont (a) survendu les capacités des algorithmes existants, et (b) créé des attentes fortes donnant lieu à une forte croissance de l'intérêt du grand public (et des étudiants) et du monde industriel.

Défis actuels. Apprentissage faiblement supervisé ; Apprentissage distribué ; Compréhension de l'apprentissage profond

Apprentissage par Renforcement

Contexte historique. L'apprentissage par renforcement (RL) a pour objet la décision séquentielle optimale, en contexte incertain ou de très grande taille. En fonction de l'état du système, on choisit une action, et on reçoit une récompense (souvent différée). L'apprentissage par renforcement et ses algo-

rithmes de base ont donné à l'IA ses premières réalisations de niveau champion du monde au XXe siècle (backgammon et jeu de dames).
Les avancées récentes reposent essentiellement sur deux approches: les bandits manchots et leurs extensions séquentielles (Monte-Carlo Tree Search, MCTS) d'une part, et l'apprentissage profond d'autre part.

- Les bandits manchots et MCTS sont un domaine d'excellence en France ; ils sont un élément crucial pour le business model du Web : l'optimisation de l'affichage des publicités sur les pages Web draine un marché considérable et fait l'objet de nombreuses start-up.
- L'apprentissage profond a obtenu en RL des réalisations impressionnantes aux jeux d'Atari (niveau humain) et de Go (niveau champion du monde et au-delà ; les avancées dans le domaine du Go sont possibles par l'hybridation des MCTS et des réseaux profonds). Ces resultats reposent sur l'entrainement de réseaux neuronaux profonds à partir de données et d'expérimentations massives, les archives des parties jouées par des humains et les millions de parties jouées contre soi-même ou d'autres IA.

Les domaines d'application où ces succès sont transposables comprennent (entre autres) les politiques d'énergie optimale et la robotique, et plus généralement les domaines qui disposent d'un simulateur assez fiable (et/ou d'un robot assez robuste) pour créer les données massives nécessaires.

Enjeux actuels. Du point de vue théorique, les garanties cherchées sont très exigeantes : on est en contexte adversarial et il suffit souvent d'une erreur pour perdre la partie. Des objectifs plus complexes (compromis gain/risque) peuvent être atteints sous des conditions restrictives.
Une autre difficulté dans un contexte d'application réelle consiste à spécifier l'objectif : il ne s'agit plus de ``gagner la partie" mais de (par ex.) ``doter un robot compagnon d'un comportement amical/approprié", ou ``d'identifier les meilleures manoeuvres de sécurité pour un réseau électrique de très grande taille".

Apprentissage profond

Contexte historique. L'apprentissage profond (à partir de 2005) n'est pas un simple revival des réseaux neuronaux : à coté des différences quantitatives (plusieurs ordres de grandeur en taille des données, en taille du modèle, et en temps calcul), une différence qualitative concerne l'importance de la représentation. La création d'une bonne représentation n'est plus une étape hors champ, laissée à la créativité des étudiants ou des experts du domaine, mais une étape centrale de l'apprentissage automatique (apprentissage « end-to-end » dans le cadre supervisé).
Un aspect essentiel des réseaux profonds est qu'ils définissent un bon espace de recherche (ils sont capables d'émuler tout programme) dans lequel il est possible de progresser (au sens d'un critère à optimiser et en fonction des données disponibles). Par contre, l'optimisation offre peu ou pas de garanties (optimisation non convexe, qualité des optima locaux, ajustement des hyper-paramètres) et demande des monceaux de données (des millions d'exemples; pour optimiser les millions de poids du réseau).
Les réseaux profonds peuvent émuler la coopération et la compétition de différents "joueurs" : par exemple un joueur génératif (capable de générer de nouvelles données) et un joueur adversarial (cherchant à distinguer les données générées des données initiales). La co-évolution de tels joueurs (réseau Generative Adversarial Network) implémente un test de Turing : il cree de nouvelles données qu'on ne sait plus distinguer des initiales. En pratique, les GANs sont à la racine de réalisations impressionnantes, mais difficiles à reproduire et d'apprentissage instable.

Transversalité. L'apprentissage profond a rapidement envahi l'ensemble des domaines de l'apprentissage . . . et au-delà.

- L'apprentissage supervisé, commen mentionné plus haut ;
- L'apprentissage non-supervisé, pour lequel la définition automatique de la représentation représente une avancée majeure ;
- L'apprentissage par renforcement, dans lequel l'apprentissage profond permet de capitaliser sur de très nombreuses réalisations du processus (voir AlphaGo).

- De manière générale, l'ensemble des domaines utilisant des simulations est en train d'être impacté.

Enjeux actuels. La pratique et la technologie sont très en avance de la théorie. Les attentes sont exubérantes et irrationnelles; en cas d'échec, la réaction consiste à augmenter la taille des données ou l'architecture des réseaux. Mais les limites ne sont pas identifiées, et du coup le domaine est en expansion avec deux dangers :
- Le battage médiatique pourrait retomber brutalement, pour un nouvel hiver de l'IA, si les promesses ne sont pas tenues (i.e., si elles sont irréalistes) ;
- Les experts, le public, les institutions s'émeuvent du coté « boîte noire » des systèmes ainsi réalisés ; pour beaucoup d'applications sensibles (de l'octroi de prêts bancaires à la recherche de terroristes), le fait de dépendre de boites noires n'est pas acceptable.

Représentation des connaissances et modélisation des raisonnements

Description de la problématique

Il s'agit de représenter les différentes informations dont a besoin un agent pour raisonner (ses croyances, connaissances, buts, etc.), au moyen typiquement de formalismes logiques, comme par exemple les logiques non monotones, les logiques para-consistantes, les logiques modales (temporelles, doxastiques, déontiques, etc), les logiques de description, etc.
Ensuite la problématique est de modéliser et d'automatiser les différents raisonnements que l'agent peut être conduit à faire, en particulier le raisonnement sur l'action et le changement des croyances, le raisonnement en présence d'exceptions, la gestion ou résolution de l'incohérence, le raisonnement temporel et spatial, le diagnostic automatisé, le raisonnement qualitatif, le raisonnement à partir de cas et par analogie, et le raisonnement sur les ontologies.

Quelques défis actuels:

- Etude de la notion de causalité: La notion de causalité a déjà été grandement étudiée en intelligence artificielle et en philosophie, mais les études actuelles ont montré la grande difficulté de définir une notion générale de causalité et de responsabilité. Il reste encore beaucoup à faire sur l'étude de cette notion, qui est de première importance pour mener à bien un grand nombre de raisonnements.

- Argumentation: L'argumentation est un processus usuel pour tenter de trouver une position acceptable à partir d'informations qui se contredisent. C'est un processus utile à un agent pour tenter de résoudre des contradictions entre ses croyances, mais c'est également un processus de base dans les échanges (délibération, persuasion, etc.) entre plusieurs agents. De plus actuellement de nombreuses plate-formes de débats, basés sur l'argumentation, se popularisent sur le web. Il est donc nécessaire de disposer de notions d'acceptation adéquates. Or, les sémantiques existantes ne sont pas forcément adaptées aux applications avec un grand nombre d'argument, et certaines relations entre argument, notamment la notion de support entre arguments, loin d'avoir une sémantique consensuellement acceptée pour le moment.

- Raisonnement sur les logiques de description: les logiques de description sont très utiles pour manipuler les informations disponibles sur le web sémantique via les ontologies. Or ces logiques sont assez pauvres, et de nombreux résultats connus pour des logiques classiques et supra-classiques ne peuvent pas être utilisés directement dans ce cadre. Il est donc nécessaire d'étudier comment mener les raisonnement correspondant dans ce cadre, pour voir si les résultats connus pour les logiques classiques peuvent s'exporter, et pour se poser la question de la calculabilité de ces méthodes dans ce cadre.

- Méthodes effectives de raisonnement (Beyond NP): De nombreuses méthodes de raisonnement (ont leur problème de décision associé ayant une complexité algorithmique qui) se situent au premier niveau de la hiérarchie polynomiale ou au delà. Or les progrès réalisés ces dernières années dans les solveurs SAT ou de contraintes, qui attaquent en pratique à présent des problèmes NP-complets de taille importante, permettent à présent d'envisager de résoudre en pratique un certain nombre de ces méthodes de raisonnement. Une initiative internationale autour de cette problématique (Beyond NP) existe, et il semble important que la France y prenne sa part, car cela pourrait aboutir à l'application pratique de ces méthodes de raisonnement.

Décision et gestion de l'incertitude

Dans les années 1970, avec les premiers travaux sur les systèmes experts, apparaît en intelligence artificielle (IA) le besoin de traiter l'incertitude, avec en particulier la nécessité de prendre en compte des règles dont la conclusion n'est pas complètement certaine (les connaissances des experts ne sont pas des théorèmes de mathématique). Les décennies suivantes voient le développement progressif de nouvelles modélisations de l'incertitude en IA, à côté des probabilités: ce sont les théories des possibilités, des fonctions de croyance, et des probabilités imprécises, par ordre croissant de généralité et de pouvoir expressif (et ordre décroissant de simplicité de mise en œuvre). Les trois théories utilisent des paires d'évaluations duales, nécessité / possibilité, croyance / plausibilité, probabililité inférieure / probabilité supérieure. Ces nouveaux cadres théoriques sont principalement motivés par le besoin de représenter des situations d'ignorance partielle, qui sont de nature épistémique (liées au manque d'information précise) plutôt qu'à la variabilité des phénomènes considérés (que capture les probabilités).

A partir des années 1980, les probabilités ont pris une place prépondérante dans la représentation de l'incertitude en IA, grâce à la masse des travaux existants sur le sujet et à l'introduction de modèles graphiques, les réseaux bayésiens, permettant de schématiser la représentation et de propager localement l'incertitude.

Des contreparties de tels modèles graphiques existent maintenant pour les trois autres cadres théoriques de représentation de l'incertitude. Mentionnons aussi les graphes conceptuels pour la représentation graphique de fragments de la logique du premier ordre en relation avec la gestion des taxonomies. Les informations fournies par les sources sont souvent incomplètes et pas forcément totalement fiables. Ceci pose la question de la synthèse et de la fusion de ces informations. Leur nature et leur qualité hétérogènes peuvent requérir des cadres théoriques permettant la représentation d'incertitudes à caractère épistémique. De même dans le raisonnement touchant au risque, il importe de bien distinguer entre incertitude épistémique et incertitude due à la variabilité. La décision (dans l'incertain, multicritère, ou de groupe) est devenue une problématique à part entière de l'IA seulement dans les années 1990. L'apport de l'IA sur ces questions est principalement marqué par un intérêt pour la représentation de préférences partielles et contextuelles, pour les modèles qualitatifs, ou pour une problématique de calcul prenant en compte la taille éventuellement considérable de l'espace des choix. Diverses extensions, quantitatives ou qualitatives, de l'utilité espérée en (décision dans l'incertain) et de la moyenne pondérée (en multicritère) ont été proposés et axiomatisés qui s'expriment en termes d'intégrale de Choquet ou d'intégrale de Sugeno (qui permettent en particulier de prendre en compte des dépendances entre critères). L'attitude du décideur, pessimiste ou optimiste, face à l'incomplétude de l'information doit aussi être prise en compte. L'idée naturelle de décision à partir de cas observés a été aussi formalisée (en termes de similarité entre cas) et justifiée. Les problèmes de décision séquentielle ont conduit l'IA à s'intéresser aux processus de décision markoviens, éventuellement partiellement observés, et à leur contrepartie qualitative.
De tout temps en IA, les cadres de représentation numériques (réseaux neuronaux, et réseaux bayésiens, notamment) tournés vers l'efficacité calculatoire et l'apprentissage ont été opposés aux cadres symboliques à base de logique et plus aptes à l'explication. Il est clair qu'un rapprochement et un interfaçage entre les deux types de cadre est souhaitable et utile dans un certain nombre de pro-

blèmes. Des travaux dans ce sens existent, certains depuis longtemps, comme la logique possibiliste (qui s'appuie sur une version qualitative de la théorie des possibilités), ou comme plus récemment la décision argumentée.

Satisfaction de contraintes et SAT (satisfaisabilité d'une formule logique)

Contexte

Une partie de la recherche en intelligence artificielle consiste à rendre la machine capable de résoudre des problèmes difficiles (pour les ordinateurs et les humains) : ce sont les problèmes dont la complexité théorique est NP-difficile.

Ces problèmes ont été étudiés dès le début de l'IA. Les jeux sont souvent NP-complets (Taquin, Sudoku, Nonogrammes, Candy Crush,. . .), voire PSPACE ou EXPTIME-complet (Othello, échecs, Go, Dames. . .). De nombreux autres domaines de l'IA soulèvent des problèmes des mêmes familles : en planification robotique, le langage STRIPS définit un problème PSPACE-complet, en apprentissage automatique, le calcul d'une réalisation "la plus probable" d'un Champ de Markov est NP-complet. La résolution de ces problèmes a un intérêt au delà de l'IA : la majorité des problèmes de vérification, conception, configuration, diagnostic, organisation dans le temps ou l'espace sont NP-difficiles (eg. la gestion des équipages en transport aérien, l'allocation spatio-temporelle optimale de cultures dans des dispositifs agro-écologiques ou la conception de protéines stables en biologie ont tous une formulation NP-difficile).

Comme l'on sait que, dans chaque famille, ces problèmes se réduisent rapidement les uns aux autres, une partie des chercheurs en IA se sont attaqués à l'élaboration de résultats théoriques et empiriques, d'algorithmes et de logiciels (solvers) capables de résoudre des problèmes difficiles énoncés dans des formalismes mathématique simples et capables de modéliser de nombreux problèmes d'intérêt (un peu comme ALICE de JL. Laurière). Les formalismes considérés sont principalement la logique propositionnelle et des modèles graphiques discrets déterministes (réseaux de contraintes) ou stochastiques (réseaux bayésiens, . . .).

Historiquement, au travers des GdR BAHIA et CPSFlex (années 90), les Français ont joué un rôle important dans ces domaines, rôle qu'ils conservent aujourd'hui (voir les répartitions des articles par pays dans la conférence CP).

Ce domaine de recherche est naturellement proche de la Recherche Opérationnelle, avec un souci particulier de facilité d'expression des problèmes symboliques (sans exclusivité). Sur les objectifs, il est également proche des techniques des recherche heuristique ou méta-heuristique.

Résultats marquants de ces dernières décennies

1994 à 2005 : la résolution d'instances aléatoires de problèmes de décision NP-complets montre l'apparition d'un phénomène de transition de phase entre problèmes avec solution et sans solution. C'est à la transition que les problèmes difficiles à résoudre en pratique abondent. Des liens s'établissent avec la physique statistique.

1990-: progrès fulgurants dans la capacité à résoudre des instance du problème NP-complet canonique SAT (trouver une valeur de variables 0/1 rendant vraie une formule logique exprimée sous forme de clauses). Le développement de logiciels Open Source, l'organisation de compétitions annuelles et l'accumulation de grands jeux d'instances industrielles y jouent un rôle important.

Certaines instances industrielles avec plus d'un million de variables booléennes sont résolues. Sur les instances aléatoires, les algorithmes heuristiques dominent (Survey Propagation). Sur les problèmes réels (industriels, jeux, scientifiques), les méthodes exactes, s'appuyant sur le ``Clause Learning'' dominent. Comme pour le "Deep Learning", ce succès reste mystérieux. Des progrès similaires sont obtenus dans le domaine des modèles graphiques déterministes (réseaux de contraintes, problème CSP), en particulier via l'introduction des``contraintes globales'' et l'exploitation explicite de la structure des problèmes. Les jeux NP-complets usuels (Sudoku, Nonogrammes, Candy Crush. . .) sont trivialement résolus par ces outils. Les Sudoku du journal "Le Monde" sont synthétisés par un ex-chercheur du domaine (Y. Georget).

2000- : ces progrès se répandent au travers de l'introduction de "langages" s'appuyant sur ces progrès fondamentaux dans les outils de résolution associés: ASP (Answer Set Programming), SMT (SAT Modulo Theory), à côté de CP. L'intérêt des industriels s'accroît car il devient possible de résoudre des problèmes industriels P-SPACE-complets (vérification de circuits digitaux, processeurs, vérification de logiciels). Intérêt similaire pour les techniques des réseaux de contraintes (CSP/CP) pour les problèmes d'organisation dans le temps et l'espace (ordonnancement, affectation complexe). Les industriels développent leurs outils, embauchent les chercheurs du domaine (eg. Laurent Perron, Google).
Les conférences sont soutenues par de grands groupes (Google, Microsoft, Intel, Cadence, IBM, Siemens,…) et des startups. Les chercheurs de l'IA mobilisent ces algorithmes pour les appliquer à des problèmes difficiles de l'IA: planification automatique (SAT-Plan),``Data mining'' (Luc de Raedt).

2000+ : Ces techniques sont étendues pour résoudre des problèmes d'optimisation (critère numérique) combinant contraintes, coûts, préférences, modèles graphiques discrets stochastiques (réseaux Bayésiens, champs de Markov) permettant de raisonner sur la base d'un modèle ``appris'' par apprentissage automatique. Il devient possible de résoudre exactement certaines instances avec 2^{1 000 000} configurations. Depuis 2010, les outils exacts ``anytime'' finissent premier de compétitions internationales d'inférence probabiliste (UAI 2010, 2012, 2014). Ces algorithmes commencent à être mobilisés pour résoudre d'autres problèmes en IA (apprentissage par renforcement, NLP, conception musicale,…) et au delà (bioinformatique, développement durable. . .). Dr Fill, entièrement écrit par M. Ginsberg, finit successivement 131e (2012), 92e (2013), 64e (2014), 41e (2016) de l'American CrossWorld Puzzle Tournament (deep learning free).

2012+: après les modèles graphiques stochastiques, introduction de contraintes globales ``Neuron'', ``Decision Tree'' et ``Random Forest'' pour élargir la gamme des modèles appris utilisables pour la construction de décisions optimisées (M. Milano/M. Lombardi/Google).

2014-: Moshe Vardi, chercheur américain en informatique théorique, président de l'Association for Computer machinery, annonce qu'il est temps de se lancer dans la résolution de problèmes au delà de NP, en particulier les problèmes de comptage (#P-complets), au cœur du raisonnement probabiliste. Les premiers algorithmes de calcul de fonction de partition avec garanties de type PAC (Probably Approximately Correct), non asymptotiques, basées sur des outils SAT apparaissent. D'autres algorithmes ``anytime'' apparaissent.

Planification et recherche heuristique

Contexte historique

Les méta-heuristiques sont des algorithmes d'optimisation basés sur une randomisation de la recherche, alternant des étapes d'exploration de l'espace de recherche et des étapes d'exploitation des meilleurs résultats passés, supervisées par une sélection des candidats-solutions elle aussi randomisée.

Comme leur nom l'indique, les méta-heuristiques sont des algorithmes d'application génériques d'heuristiques spécifiques au domaine et à l'espace de recherche, se rapprochant en cela de la plupart des méthodes exactes, dans le cas par exemple de l'optimisation combinatoire (e.g., choix de l'ordre de parcours des variables dans la programmation par contraintes, croisement de solutions pour les problèmes de voyageur de commerce, etc). Elles diffèrent entre elles par les différents agendas d'application des opérateurs de recherche dans l'espace des solutions qu'elles préconisent. À l'inverse des méthodes exactes, elles n'offrent aucune garantie, mais permettent par contre de s'attaquer à des problèmes peu structurés qui échappent aux approches exactes, comme on en rencontre fréquemment dans le monde réel.

Ainsi, l'un des algorithmes les plus performant dans le domaine de l'optimisation continue est l'algorithme CMA-ES (Covariance Matrix Adaptation Evolution Strategy), algorithme évolutionnaire à base de transformations gaussiennes – dont on vient d'ailleurs de prouver qu'il n'est finalement rien d'autre qu'un algorithme de gradient naturel dans l'espace des distributions gaussiennes. Et certains ``records'' pour des problèmes du voyageur de commerce de très grande taille sont détenus (ou l'étaient encore il y a peu) par un algorithme utilisant des croisements de solutions très spécifiques.

Mais l'on retrouve ici un problème aujourd'hui essentiel en optimisation (comme d'ailleurs en apprentissage), celui du réglage des hyper-paramètres : on sait bien qu'il n'existera pas d'algorithme universel, et chaque instance de problème est résolu de manière optimale (la plus rapide pour atteindre une solution, ou donnant la meilleure solution en un temps compté) par un algorithme particulier réglé de manière spécifique. On parle aujourd'hui de Programming by Optimization (après H. Hoos), ce qui peut aller, pour une classe de problèmes donnée, du choix du meilleur algorithme parmi un portfolio (et on rejoint alors le domaine de la recommandation) jusqu'à l'optimisation d'heuristiques pour des algorithmes exacts, voire la conception d'une hyper-heuristique, algorithme construit à partir de briques de bases algorithmiques.

Par exemple, le niveau zéro de l'approche – la simple optimisation des hyper-paramètres d'algorithmes existant – a pu apporter des accélérations de plus d'un ordre de grandeur sur un ensemble de solveurs, qu'ils soient spécifiques (comme dans le domaine du planning) ou génériques (comme le produit commercial CPlex).

Enjeux actuels

Le corpus de résultats théoriques concernant la convergence des méta-heuristiques s'accroissent rapidement, mais reste encore bien en deçà des succès pratiques. Et de toute façon, ils ne pourront concerner que des résultats de complexité (en moyenne, ou dans le pire cas), les garanties d'optimalité devant être recherchées ailleurs.

L'application pratique du concept de Programming by Optimization pose plusieurs problèmes: celui de la représentation des instances de problème, via un ensemble de descripteurs, qui sera l'espace de recherche de la méta-optimisation ; celui du choix de la classe d'instances pour lesquelles on va chercher des hyper-paramètres optimaux ; et celui de l'algorithme d'optimisation ou d'apprentissage qui sera utilisé pour apprendre les hyper-paramètres correspondant à un ensemble d'instances donné.

Agents autonomes et systèmes multi-agents

Cette thématique se concentre sur l'analyse, la conception et l'implémentation de systèmes composés d'entités autonomes et/ou en interaction. La thématique ``Agents Autonomes et Systèmes Multi-Agents'' est la cible de la conférence AAMAS et de la revue Journal of Autonomous Agents and Multi-Agent Systems. Une partie de cette thématique est transversale en ceci qu'elle rejoint d'autres sous-branches de l'IA, notamment la représentation des connaissances et le raisonnement pour les problématiques liés aux sociétés d'agents (représentation des états cognitifs des agents, argumentation, né-

gociation...); la planification pour la réalisation de tâches complexes par des agents autonomes ou des groupes d'agents en collaboration; l'apprentissage multi-agent et l'émergence de comportements collectifs. D'autres sont propres au domaine, notamment:

1. la conception de langages et de protocoles pour la communication, le dialogue et la collaboration entre agents, leur permettant, entre autres, de former des coalitions.

2. la résolution collective de problèmes : il s'agit de résoudre de manière décentralisée un problème qui se pose globalement à la collectivité d'agents (et donc aux humains qui les ont programmés). Un exemple typique d'application : résolution d'une tâche d'exploration dans un environnement mal connu par une flotte de robots se déplaçant de façon autonome et pouvant (plus ou moins bien) communiquer entre eux. On est ici en plein dans le cadre de l'IA distribuée, avec des intersections fortes avec les domaines de la planification, du raisonnement et de la décision dans l'incertain, voire de la robotique.

3. l'étude de sociétés d'agents et des approches sociétales des systèmes multi-agents (organisations et institutions, systèmes socio-techniques, systèmes normatifs, confiance et réputation), ainsi que les problèmes normatifs et éthiques qui en découlent (par exemple : évitement et gestion d'accidents relatifs à des voitures autonomes en interaction dans une même zone de conduite).

4. l'étude de modèles représentant le comportement d'agents en interaction, provenant entre autres de la microéconomie (notamment la théorie des jeux, coopératifs ou non-coopératifs) et visant à mesurer l'impact des comportements stratégiques sur l'efficacité globale du système, et à concevoir des mécanismes permettant d'éviter ces comportement stratégiques ou de limiter leur interaction. On est ici dans le domaine de la théorie de la décision algorithmique, qui a sa conférence dédiée (Symposium on Algorithmic Game Theory) et qui est l'une des thématiques de la conférence ACM Economics and Computation. Les applications concernent notamment les enchères (en particulier les enchères combinatoires) et, au-delà, le commerce électronique, ainsi que les problèmes de routage distribué (jeux de congestion), et les systèmes P2P.

5. la conception et l'étude de mécanismes pour la décision collective (vote, partage de ressources, appariement, formation de groupes). On est ici dans le cadre du choix social computationnel, qui a son workshop dédié (COMSOC). Il s'agit à la fois d'utiliser des techniques de l'IA pour résoudre des problèmes de décision collectives algorithmiquement complexes, et d'appliquer des concepts, des modèles et des techniques proviennent du choix social pour la gestion de collectifs d'agents (humains et/ou artificiels).

6. la modélisation et la simulation de phénomènes complexes. Il s'agit ici d'utiliser les systèmes multi-agents pour simuler des interactions existant entre agents autonomes, et de chercher à déterminer l'évolution de ce système afin de prévoir l'organisation qui en résulte. Ce paradigme permet même d'expérimenter des scénarios qui ne seraient pas réalisables sur des populations réelles, que ce soit pour des raisons techniques ou éthiques. Ce qui importe, c'est le comportement d'ensemble, non pas le comportement individuel. Des exemples concrets : simulation de la propagation de maladies dans un espace géographique paramétrable, simulations de trafic routier ou d'évacuation de foule paniquée dans un bâtiment, qui permettent d'étudier l'origine des problèmes (embouteillages, bousculades) afin de réfléchir à des solutions pour les restreindre.

IA & Web

Le Web sémantique est un domaine d'application emblématique de la thématique de la représentation des connaissances et du raisonnement au service du Big data. La représentation des connaissances et du raisonnement se caractérisent par la représentation formelle de connaissances et leur traitement symbolique par des algorithmes de raisonnement automatique. Dans le cadre du Web sémantique, les

connaissances sont formalisées sous forme d'ontologies qui permettent de typer et relier sémantiquement les données du Web pour les interroger de manière plus pertinente et plus efficace. Une ontologie est un vocabulaire structuré de classes et de propriétés définies à l'aide d'un formalisme logique compréhensible par des humains et traitable par des machines. L'algorithmique à mettre en place pour répondre à des requêtes posées par le biais d'ontologies repose sur des algorithmes de raisonnement automatique permettant d'inférer, à partir des connaissances formalisées dans l'ontologie, que des éléments de réponses trouvées dans les données satisfont bien la requête. Compte tenu de la taille très importante des données, garantir des temps de réponses raisonnables pour l'évaluation des requêtes est un problème central au cœur des travaux sur le Web sémantique. Il s'agit de trouver des restrictions sur l'expressivité du langage de requêtes et du langage d'expression des ontologies permettant de garantir une complexité algorithmique du calcul des réponses qui soit inférieure dans tous les cas à un polynôme de la taille des données.

Au-delà du Web sémantique, les ontologies servent à lier et intégrer des données multiples et hétérogènes dans de nombreux autres domaines. Elles servent aussi d'interfaces ``intelligentes'' de requêtes permettant aux utilisateurs d'exprimer des requêtes à un niveau d'abstraction adapté à leurs besoins et à leurs compétences. La création d'ontologies est un problème difficile souvent abordé dans un mode semi-automatique : un noyau est d'abord défini manuellement par un expert du domaine concerné par les données à intégrer et analyser, puis est ensuite enrichi par fouille de données.

Parce qu'il est symbolique, le raisonnement sur des données décrites par des ontologies est transparent et traçable. Un résultat d'analyse ou une réponse à une requête sont explicables : la logique conduisant à l'obtention du résultat est explicite, les sources et liens peuvent être remontés, vérifiés et interrogés et les résultats sont ainsi mieux appréhendés par un utilisateur non informaticien.

Traitement automatique des langues

Le traitement automatique des langues naturelles (TALN) est une discipline de l'informatique et des sciences du langage qui s'intéresse à la modélisation et l'automatisation des processus cognitifs langagiers, que ce soit la compréhension de messages, la lecture de textes, le dialogue, la traduction, l'acquisition de connaissances à partir de textes, etc.

Le TALN définit des algorithmes, construit des ressources langagières, élabore des outils et des architectures logicielles pour le traitement des langues. Les algorithmes soit s'ancrent sur les données et relèvent de l'apprentissage automatique - traitement empirique ou statistique de la langue, soit modélisent les processus langagiers - traitement symbolique de la langue.
Les outils et les ressources développées servent soit à assister les humains dans leur travail (correction orthographique, traduction ou dictée automatique) ou soit pour opérer à une grande échelle inaccessible à l'humain (moteurs de recherche, fouille de textes et analyse de contenus).
L'approche empirique qui a émergé dans les années 1990 est dominante du aux succès qu'elles a engrangés ces dernières années pour des applications grand-public tel que la traduction automatique, la reconnaissance de la parole ou encore la détection des opinions sur les réseaux sociaux.

La communauté internationale du TALN est importante. Elle est structurée autour de l'ACL (Association for Computational Linguistics) qui organise ses conférences, annuelles : ACL (CORE A*), 1 200 participants et 900 articles soumis en 2016, NAACL (CORE A), EMNLP (CORE A), et bi-annuelle : COLING (CORE A) 1 200 participants, EACL (CORE A) chapitre européen de ACL, 500 participants, IJCNLP chapitre asiatique (CORE B), 800 participants, etc. La communauté française du TALN est structurée autour de son association savante, l'ATALA, depuis 1959. L'ATALA organise deux conférences annuelles, TALN et RECITAL (pour les doctorants) et des journées d'étude. Elle édite la revue TAL en langue française. Une réflexion pour la structuration scientifique de la communauté sous l'égide de l'INS2I est en cours.

Robotique

La problématique de l'intelligence artificielle telle que posée par Alan Turing (A. M. Turing (1950) Computing Machinery and Intelligence. Mind 49: 433-460.) était de savoir si les ordinateurs pouvaient être capables de « pensée » (Can Machines Think?) et il l'a immédiatement traduite par la question de l'imitation de l'homme. Il faut bien noter que la question initiale de Turing portant sur la « pensée » a été traduite par les fondateurs de l'IA par celle de l' « intelligence », ou des « mécanismes de haut niveau », ce qui avait eu pour conséquence de casser le lien entre le raisonnement, l'action, la perception, l'apprentissage et l'interaction.

La problématique de la robotique pose l'ensemble de ces questions. Le robot-machine est soumis à la complexité du monde réel dans lequel il évolue et il doit respecter sa dynamique. La notion d'intelligence doit alors être posée de manière à rendre compte globalement des processus sensorimoteurs, perceptuels et décisionnels permettant l'interaction en temps-réel avec le monde, en tenant compte des contraintes d'incomplétude et d'incertitude de perception ou d'action. Les cinq fonctions principales d'un robot sont : la perception/interprétation, le mouvement/action, le raisonnement/planification, l'apprentissage, et l'interaction avec d'autres agents. Il est nécessaire d'adopter une vision d'ensemble du robot, en tant que système intégrant ses différentes capacités et permettant à la fois la réactivité et la prise de décision sur le long terme. Ces fonctions doivent être intégrées de manière cohérente dans une architecture de contrôle globale (architecture cognitive).

Perception, action, apprentissage

De nombreuses avancées ont été réalisées dans chacune des fonctions fondamentales du robot. Dans les années 1985-2000, la problématique de la localisation et de la cartographie simultanées (SLAM) a connu un développement formidable qui a permis de bien en cerner les fondements et de produire des systèmes efficaces. Le point faible important restant le manque d'interprétations plus sémantiques de l'environnement et des objets. L'apprentissage profond a fait une incursion considérable en robotique, prenant la place de l'apprentissage bayésien, à la fois pour la perception et pour la synthèse d'actions.
Mais la perception en robotique nécessite une interaction du robot avec son environnement et non une simple observation. Le lien entre perception et apprentissage, en particulier apprentissage par renforcement est essentiel pour extraire la notion d'affordance, qui rend compte des propriétés perceptuelles des objets, qui va jusqu'aux représentations symboliques, mais qui permet aussi d'apprendre des compétences (skills) ou des capacité d'action pour le robot.

Interaction

Les problèmes d'interaction et d'action conjointe homme-robot, avec la mise en œuvre de prise de perspective, est une problématique fondamentale pour permettre une interaction efficace et naturelle entre l'humain et le robot. Ce sujet demande un développement qui associe des recherches en robotique et en SHS (en particulier, en sociologie, philosophie, linguistique). Le rôle des émotions dans l'interaction est à explorer, bien au delà de travaux actuels qui se contentent de classer a priori des expressions faciales ou de produire des expressions d'émotions artificielles par le robot.

Décision

La prise de décision s'appuie sur des techniques relativement classiques de l'IA. La non-adéquation de la planification déterministe avec les contraintes du monde réel a amené le développement de méthodes basées sur les processus décisionnels dans l'incertain, comme les processus markoviens. Ces approches trouvent leur limites dans une trop grande complexité. L'apprentissage d'actions en robotique est souvent réalisé par des méthodes d'apprentissage par renforcement et le lien entre planifica-

tion et apprentissage n'est pas très clair. Des recherches sur de nouvelles approches alliant apprentissage et raisonnement probabiliste avec une convergence rapide sont nécessaires.

Conscience de soi

Mais la question initiale de Turing ne devrait-elle pas conduire à l'interrogation suivante : une machine peut-elle avoir une faculté de conscience d'elle-même ? Car malgré toutes les recherches en robotique et IA, les résultats, aussi significatifs soient-ils, restent le plus souvent applicables dans des contextes restreints et bien définis. Ainsi, la perception ne permet pas à un robot de comprendre son environnement, c'est à dire d'élaborer une connaissance suffisamment générale et opératoire sur celui-ci, la prise de décision reste limitée à des problèmes relativement simples et bien modélisés. Les principes fondamentaux restent largement incompris, qui permettraient aux robots d'interpréter leur environnement, de comprendre leurs propres actions et leurs effets, de prendre des initiatives, d'exhiber des comportements exploratoires de curiosité, et d'acquérir de nouvelles connaissances et de nouvelles capacités. Les clés pour permettre la réalisation de ces fonctions cognitives peuvent être le méta-raisonnement et la capacité d'auto-évaluation, deux mécanismes réflexifs.

Pluridisciplinarité

La recherche en robotique pose des questions proches des sciences cognitives, des neurosciences et de plusieurs domaines des SHS, comme la sociologie et la philosophie. Des programmes interdisciplinaires seront probablement la bon moyen pour aborder les différentes facettes des questions fondamentales.

Vision et reconnaissance des formes

La recherche en vision artificielle connaît des développements sans précédent depuis une quinzaine d'années sur des aspects en lien avec l'intelligence artificielle, à savoir la reconnaissance des formes, la détection d'objets, la classification et l'interprétation d'images. A cela s'ajoute désormais l'apprentissage de représentations, de métriques pour la comparaison et la recherche d'images.

La vision est un domaine phare pour les méthodes d'apprentissage. A titre d'illustration, les trois exemples remarquables cités dans le livre blanc sur l'IA publié en 2016 par l'INRIA sont dans ce domaine. Cette réussite repose à la fois sur des avancées en termes de modélisation des représentations de l'information visuelle, sur des méthodes d'apprentissage statistiques exploitant des données massives annotées, et enfin sur des calculateurs très puissants adaptés aux types de calcul nécessaires (GPU).

Tout comme le TAL et l'Audio, la Vision est fortement impactée par l'apprentissage profond (deep learning). Les approches de modélisation reposant sur des architectures neuronales profondes, en particulier les réseaux de neurones convolutionnels sont aujourd'hui les approches dominantes. Bien que s'appuyant sur des principes introduits il y a une trentaine d'années, en lien avec des résultats de travaux en neurobiologie sur la vision animale, il y a eu des avancées récentes sur la compréhension des structures de ces réseaux et de leur apprentissage.

Si ces méthodes, en particulier par apprentissage profond, ont beaucoup de succès dans les compétitions de reconnaissance visuelle, plusieurs thèmes de recherche actuels constituent des défis importants :
• Segmentation sémantique et détection à large échelle
• Vision et Langage (du mot clé à la narration)
• Modélisation des architectures de vision (invariance, structure, compression)
• Composition (contextualisation des détections, apprentissage séquentiel)
• Apprentissage (très) faiblement supervisé, avec des données bruitées, avec des données synthétiques

• Apprentissage non supervisé, génération d'images (modèles génératifs, GAN)
• Multi-taches/adaptation/transfert (classification, détection, segmentation, ...)
• Modèles d'attention visuelle
• Intégration d'informations multimodales (images + texte, son, etc.)
• Exploitation de connaissances a priori
• Représentation des données (géométriques, multivariées, creuses, par dictionnaires, structurées, etc.)
• Explicabilité des résultats
• Constitution de bases d'apprentissage suffisamment importantes (intégrant des spécificités liées au domaine pour la télédétection, l'imagerie médicale, …)
• Génération automatique d'exemples d'apprentissage
• Images fixes vs. vidéos et 3D+t (reconnaissance d'actions par exemple)

Les enjeux en recherche sont de taille car dès lors que l'on peut doter une machine d'un œil performant (dans un domaine plus ou moins ciblé), sans parler de révolution cambrienne, les applications industrielles sont très importantes. Au-delà de la reconnaissance, ces méthodes sont maintenant exploitées pour des tâches de plus bas niveau, comme la restauration, le recalage, etc. Le domaine récent de la co-conception (capteurs + algorithmes) devrait aussi en bénéficier.

Les méthodes inspirées de "l'intelligence humaine" ont donné lieu au développement de méthodes et algorithmes souvent cantonnés aux tâches de bas niveau pendant longtemps. Pourtant l'importance des approches linguistiques et de la sémantique a été identifiée dès les années 1960. Ces approches connaissent aujourd'hui un regain d'intérêt et méritent d'être développées, en allant bien au-delà des anciens systèmes experts qui avaient été explorés il y a un temps pour l'interprétation d'images. Les approches associant représentation des connaissances, raisonnement et vision sont donc également un défi. Ce sujet difficile mais crucial devrait donner lieu à des recherches dans les prochaines années. Des résultats en termes d'interprétabilité devraient également en découler. Notons enfin les interactions fortes de ce thème avec ceux de la représentation des connaissances et du raisonnement, de la décision et de la gestion de l'incertitude (pour la modélisation et la prise en compte des incertitudes inhérentes aux images et aux connaissances, et pour les méthodes de décision), de l'apprentissage, des ontologies, de la robotique, et potentiellement presque tous les autres thèmes pour les méthodes qui y sont développées.

Modélisation cognitive et systèmes neuro-informationnels

L'Intelligence Artificielle (IA) est le nom donné à l'ensemble des techniques qui visent à reproduire dans des machines électroniques les aptitudes sensorielles et cognitives des animaux, de l'homme en particulier. L'intelligence des êtres vivants est multiple, variée et évolutive. Il en est de même de son avatar technologique : l'Intelligence Artificielle (IA) qu'il n'est donc pas possible de définir au singulier.

L'intelligence peut toutefois se définir selon deux grandes catégories de facultés. La première intelligence, que nous partageons avec la plupart des espèces animales, est celle qui nous relie au monde extérieur et nous permet de percevoir, d'apprendre, de reconnaître, d'estimer et de décider. Elle est centrale à notre capacité d'adaptation et de survie dans un monde toujours fluctuant et son automatisation a été jusqu'ici le principal moteur de l'intelligence artificielle (IA). Grâce aux progrès constants de la microélectronique, de l'informatique, du traitement du signal, de l'analyse statistique et plus récemment de l'apprentissage profond opérant sur de vastes données (big data), des résultats remarquables ont été obtenus dans l'automatisation des tâches de perception et de décision.

La seconde intelligence est plus intime. Elle englobe les facultés de l'esprit, celles qui nous permettent d'imaginer, d'élaborer, d'inventer et d'espérer. Le seul modèle dont nous disposons pour essayer de reproduire dans une machine les propriétés de cette intelligence créatrice est notre cerveau, dont l'architecture est nettement différente de celle de l'ordinateur classique (mémoire et processus s'y entre-

mêlent dans de mêmes circuits, aujourd'hui encore très mystérieux). Une nouvelle informatique, nécessairement neuro-inspirée, est à imaginer pour reproduire cette seconde intelligence, autrement dit faire émerger ce qu'on appelle l'IA générale ou l'IA forte.

Le système nerveux est connexionniste, c'est-à-dire que son fonctionnement repose sur des transferts de signaux entre les composantes (neurones ou groupes de neurones) d'un vaste réseau préorganisé par la neurogenèse. Une question fondamentale porte sur le degré minimum de fidélité physicochimique qu'il est nécessaire d'adopter pour comprendre l'information mentale, en particulier la façon dont elle est mémorisée et remémorée. Robustesse et pérennité sont, sur cette question, des propriétés de première importance. Une distinction doit notamment être faite entre les circuits périphériques du système nerveux et les circuits corticaux. Dans le premier cas, il s'agit principalement de canaux nerveux unidirectionnels (feedforward) dans lesquels instants d'arrivée et fréquence des potentiels d'action (spikes) sont déterminants. Dans le second cas, celui de circuits massivement bouclés, il est possible d'imaginer que c'est le graphe des connexions qui est porteur de l'information et non pas les caractéristiques des signaux en eux-mêmes. En particulier, la robustesse de l'information mentale peut être expliquée par la forte redondance des connexions entre les sommets d'un même sous-graphe représentatif d'un élément de connaissance.

Cette hypothèse d'information corticale graphique, la cognition se développant et se maintenant donc sur un réseau numérique, ouvre de belles perspectives dans l'émulation informatique du fonctionnement cortical et par conséquent dans les progrès de l'IA générale. Il est étonnant d'observer que les rapports récents sur les stratégies à adopter pour le développement de l'IA, notamment aux Etats-Unis, ne préconisent pas de rassembler neuroscientifiques et informaticiens dans de mêmes équipes de recherche, condition indispensable à la compréhension de la cognition et à son émulation dans des circuits électroniques. Tout se passe comme si, dans les différents groupes de réflexion, l'Intelligence Artificielle n'avait pas grand-chose à voir avec l'Intelligence Naturelle.

Il va de soi qu'une meilleure compréhension du cortex cérébral, à l'échelle informationnelle, non seulement bénéficiera au développement de l'IA mais contribuera également aux progrès des neurosciences cognitives. Il serait par exemple possible de réaliser des simulations in silico pour essayer de mieux comprendre certains troubles mentaux et éventuellement de trouver de nouvelles thérapies.

Interaction Humain-Agent

L'intelligence artificielle a pris, ces dernières années, une place de plus en plus importante dans le domaine de l'interaction humain-machine, conduisant à de nouveaux paradigmes d'interaction. De nombreux travaux de recherche portent sur l'adaptation des interfaces aux utilisateurs, la prise en compte de leurs individualités mais aussi de leurs états affectifs, de leurs capacités cognitives et sensori-motrices, de leurs intentions, de leur culture… L'interface doit donc s'adapter dynamiquement aux besoins et états affectifs des utilisateurs, en tant qu'individu, collectif ou sociétal, pour améliorer non seulement les performances de l'utilisateur mais aussi la qualité de leur expérience (déterminée par leur engagement, leur degré de confiance, etc). Dans cet objectif, les modes d'interaction font maintenant usage de plusieurs modalités telles que la voix, le comportement non-verbal, la prosodie, le toucher. Les interactions humain-machine demandent de prendre en compte les signaux émis par l'humain (ce qu'il dit, comment il le dit, ce qu'il fait..), de générer une réponse multimodale, de définir des stratégies d'interaction. La même information doit être accessible à tout moment sur des dispositifs miniaturisés ou au contraire sur de très grandes surfaces, tout en conservant une interaction « naturelle » intégrée dans l'environnement. De plus en plus d'objets connectés permettent de capturer des données physiologiques sur les utilisateurs et, dans le même temps, les écrans sont partout, sur toutes les surfaces. Enfin, les interfaces avec une représentation humanoïde (robot physique ou virtuel - souvent appelé agent conversationnel animé) tendent à se généraliser. Elles doivent être dotés de capacités

proches de celles des humains : percevoir, comprendre les intentions, raisonner, prendre des décisions, montrer des affects, des intentions, etc.

Plusieurs défis sont identifiés par la communauté scientifique :

Reconnaissance : Reconnaissance de la parole multimodale, états affectifs, états mentaux, actions et intentions à partir de signaux multimodaux ; reconnaissance du contexte ; fusion des signaux (les 5 sens : audio, visuel, toucher, goût et odeur ; signaux physiologiques, objets connectés)
Autonomie et Adaptation : Adaptation dynamique de l'interface à l'environnement, aux besoins et aux tâches des utilisateurs ; adaptation dynamique aux états affectifs et mentaux des utilisateurs ; Augmenter l'autonomie des agents conversationnels physiques et virtuels
Prises de décisions : Prédire les actions de l'utilisateur, aider à la prise de décision (côté entrée) ; générer des prises de décision et montrer des comportements multimodaux (côté sortie)
Dialogue : Dépasser les échanges courts d'une interaction en maintenant l'engagement de l'utilisateur (passer d'un modèle de QA à un modèle de dialogue)
Systèmes de formation : doter les agents pédagogiques d'une représentation des connaissances pour expliquer leur raisonnement et actions
Évaluation : Être crédible, donner confiance, ne pas tomber dans la vallée de l'étrange, développer des techniques d'évaluation, des benchmarks
Visualisation : visualiser et interagir une grande quantité d'informations sur des dispositifs hétérogènes et distribués (eg miniaturisés, très grands et déformables)
Explicabilité : visualiser pour comprendre le processus des algorithmes d'apprentissage
Éthiques : Respecter la vie privée des utilisateurs et des règles éthiques (eg regardant la simulation des émotions)

Les méthodes d'Intelligence Artificielle employées dans ces nouvelles IHM reposent à la fois sur des modèles experts (par exemple écrite dans des logiques formelles) et sur l'analyse automatique (par exemple à l'aide d'algorithme d'apprentissage) de données multimodales d'interaction humain-humain ou humain-machine. Il faut donc développer des outils d'annotation de ces données, d'analyse et de synthèse de comportement, de dialogue, de voix de synthèse.

Annexe 6 – Rapport du sous-groupe dédié à l'IA pour le droit

La fertilisation croisée du droit et de l'IA est indéniable et peut conduire à des avancées sociétales majeures. L'interaction de l'IA et du droit se situe non seulement dans le partage d'informations mais aussi dans la production de services juridiques ou d'aides à la décision.

Le développement récent des *legaltechs* souligne l'importance croissante et la dynamique particulièrement prometteuse du secteur. Leur développement est particulièrement dynamique. Ainsi, le marché des *legaltechs* aux États-Unis a notamment levé depuis 2011 environ 740 millions de dollars, pour un marché du droit dont le volume est de l'ordre de 400 milliards de dollars (« Tech in the Courtroom : Legal Tech Companies See Lackluster Funding », *CBS Insights*, 01/072016 ; « The size of the US Legal Market Shrinking Piec or a bigger Pie-an LEI graphic », *Legal Executive Institute*, 11/01/2016).

L'Union Européenne a également investi dans ce secteur. Désormais, une plateforme permet de saisir en ligne, dans le cadre d'un litige relatif à la consommation, les éléments de nature à résoudre un litige entre un professionnel qui vend des biens ou des services en ligne dans l'UE et un consommateur, sans passer par un tribunal. Ce dispositif suppose néanmoins l'intervention de personnes qui prennent contact avec le professionnel visé et le consommateur concerné et va permettre de parvenir à la solution du litige.

Toutes ces évolutions s'inscrivent aujourd'hui dans le cadre de la justice prédictive. Une étude publiée en 2017 dresse (*Droit & Digital : Réalité et prospective*, Day One 2017) sur la base de l'analyse réalisée dans 12 pays et sur quatre continents, la cartographie de 140 *legaltechs* qui montre que celles-ci se positionnent aujourd'hui essentiellement dans trois secteurs :
- les legaltechs développant du partage d'informations (35%),
- celles visant à la production de services juridiques (54%)
- et enfin celles relatives au développement de services d'aide à la décision (11%).

Ces legaltechs permettent de générer automatiquement des documents juridiques (contrat, mise en demeure) ou judiciaires (assignation). Elles permettent également d'obtenir des statistiques sur les décisions des tribunaux, comme par exemple le montant des indemnisations allouées et de dégager ainsi la position d'une juridiction déterminée, voire très précisément d'un magistrat donné.

Mais, il n'existe à l'heure actuelle aucune IA qui permettrait d'aller plus loin et de permettre de résoudre de manière automatisée le différend voire de permettre d'intervenir en quelque sorte comme « arbitre » dans la résolution du conflit qui oppose les parties prenantes.

En droit, la solution d'un litige est fondée sur une application stricte des lois et une utilisation pertinente des jurisprudences. Il en est ainsi dans le cadre des conflits entre personnes, les divorces, les conflits dans les relations de travail, les accidents médicaux ou de la circulation. Dans toutes ces situations le juge rend sa décision en jouant le rôle « d'arbitre » quant à la résolution du conflit qui oppose les parties prenantes.

Plusieurs développements importants en Intelligence Artificielle (IA) concernent des techniques de représentation des connaissances et de modélisation du raisonnement (p.ex. argumentation computationnelle, raisonnement par analogie, raisonnement abductif, raisonnement inductif, etc.) d'entités computationnelles (appelées « agents intelligents »). Par ailleurs des développements importants concernent des techniques d'interaction et de dialogues automatisés entre plusieurs entités computationnelles. Cet ensemble de techniques a permis la modélisation des conflits et l'automatisation de la prise de décisions pour leur résolution.

Des questions sur l'utilisation de l'intelligence artificielle pour résoudre de manière automatisée (ou semi-automatisée) un certain nombre des problèmes rencontrés en droit se posent légitimement. De telles questions pourraient être par exemple :

1) Quelle aide l'IA peut-elle produire pour identifier les éléments pertinents de la jurisprudence ?

L'intelligence artificielle propose déjà une aide significative dans le domaine du droit. En effet plusieurs start-ups en France (p.ex., Doctrine, Legalife, Legalstart, Captain Contrat, Predictice, etc.) et aux USA (Rocket Lawyer, Legalzoom, eBrevia, NextLaw labs, Lex Machina, LexPredic, etc.) proposent différents types de services allant de la génération des documents juridiques ou judiciaires en prenant en compte des informations fournies par les utilisateurs jusqu'au calcul des probabilités de résolution d'un litige, le montant probable des indemnités, l'analyse des contrats, des statistiques sur les performances des avocats ou le profilage des juges en détectant les décisions prises habituellement par ces juges dans différentes situations, etc. En France l'initiative « Open Law » concernant l'ouverture des données juridiques mais aussi la même initiative au niveau européen avec Open Law Europa, ont comme objectif de faciliter le développement des services décrits ci-dessus. Ces progrès peuvent augmenter les chances de succès pour les justiciables et s'adressent à la fois aux justiciables et au monde juridique (p.ex., les avocats).

Tous ces services sont essentiellement fondés sur l'utilisation des techniques d'analyse de données qui exploitent des larges bases avec des jurisprudences et autres informations concernant des cas antérieurs. Dans ce contexte il y a une catégorie spécifique d'algorithmes développés au cours des dernières années pour des tâches prédictives comme celles évoquées ci-dessus. Ces algorithmes sont appelés algorithmes prédictifs. Aux USA ces algorithmes sont essentiellement utilisés dans des cas comme l'estimation de la caution avant le procès, la détermination de la peine criminelle, la liberté conditionnelle, la justice juvénile.

2) L'IA peut-elle automatiser la prise de décision juridique dans des cas où la part de subjectivité est limitée ?

Il existe une quantité considérable de travaux théoriques concernant la représentation des connaissances et le raisonnement dans le contexte juridique, en utilisant différents types de logiques. Cependant, jusqu'à présent, il existe un nombre de travaux très limité sur les systèmes juridiques qui sont capables de simuler le raisonnement des juges leur permettant de prendre des décisions indépendamment du domaine d'application de la loi. En effet, les systèmes existants (voir ASHSD-II, CHIRON, JUDGE, Split-Up, SHYSTER, TAXMAN) ne traitent pas dans sa totalité la complexité du raisonnement des juges et modélisent des tâches très spécifiques dans des domaines particuliers au sein des systèmes de juridiction des différents pays (par exemple les litiges relatifs aux biens matrimoniaux en droit anglais, les activités de planification fiscale en vertu de la législation et des codes fiscaux des États-Unis, la répartition des biens matrimoniaux conformément à la loi australienne sur le droit de la famille, etc.). Pour autant, la possibilité d'un mécanisme de raisonnement juridique générique (intégrant les différents types de raisonnement utilisés par les juges lors de la prise de leurs décisions) pour développer des systèmes juridiques (fondés sur l'IA) génériques pouvant être instanciés afin de traiter différentes situations d'application du droit, est une perspective très prometteuse qui ouvre un nouvel horizon concernant la contribution de l'intelligence artificielle au domaine juridique.

L'intelligence artificielle (IA) est maintenant capable de mener à bien de nombreuses tâches sophistiquées, y compris le soutien de certains aspects de la prise de décision juridique, en prenant en compte la loi ou la jurisprudence. Ces tâches sont toutefois largement soutenues de manière isolée par des systèmes qui sont généralement statiques et spécifiques à un domaine précis (comme c'est le cas des systèmes évoqués ci-dessus). En outre, ces systèmes fonctionnent souvent comme des « boîtes noires », capables de fournir une recommandation sans ou avec une justification mais limitée. Or, un objectif supplémentaire pourrait être de fournir des systèmes unifiés basés sur l'intelligence artificielle, que nous appelons «bots juridiques», pour soutenir la prise des décisions juridiques en général, en tenant compte des lois/règlements et des cas passés (jurisprudences) en évolution, et en mettant les décideurs (p.ex. les juges) au centre du système. En particulier, nous envisageons des systèmes qui évoluent tout au long de leur cycle de vie, capables de tenir compte des changements dans les législations/réglementations ainsi que des bases de cas antérieurs (jurisprudences) croissant dans le temps.

Ces systèmes seront capables d'améliorer leurs recommandations dans le temps en interagissant avec les utilisateurs, en justifiant leurs recommandations et en tenant compte des commentaires des utilisateurs (les juges). La condition préalable à la réalisation de cet objectif est la pluridisciplinarité, qui comprend des experts juridiques, des analystes de données et des chercheurs en IA dans le domaine de la représentation des connaissances et du raisonnement, l'apprentissage automatique et le traitement du langage naturel.

Un « bot juridique » est le résultat de l'intégration de (1) un mécanisme de raisonnement utilisant la loi/réglementation et (2) un mécanisme de raisonnement utilisant les cas antérieurs (jurisprudences). Les informations utilisées en (1) et (2) sont extraites de façon semi-automatique des textes stockés dans des bases de données publiques, décrivant les lois et les jurisprudences et constituent la troisième composante (3) construisant les bots juridiques. Plus précisément pour (1), il conviendrait d'utiliser un raisonnement fondé sur l'argumentation avec des règles et des préférences sur elles alors que pour (2) il faut utiliser un raisonnement fondé sur l'argumentation pour analyser les conflits entre les cas antérieurs (quand ils sont tous similaires au nouveau cas mais ont donné lieu à des résultats/décisions différents). L'information en (3) sera obtenue au moyen des techniques d'analyse des données et du traitement du langage naturel pour extraire les structures d'arguments des cas antérieurs (jurisprudences) et des techniques d'exploration d'arguments (argument mining) pour extraire la structure des arguments des cas antérieurs. Aussi, il conviendrait d'extraire des règles et des préférences à partir d'un raisonnement par cas, via son incarnation basée sur l'argumentation (l'intuition étant que les décisions sur les cas antérieurs auraient dû suivre certaines règles et avoir pris en compte des préférences). La combinaison de ces techniques pourrait fournir des suggestions justifiées basées sur l'argumentation (par exemple, des prédictions et des recommandations sur de nouveaux cas). L'utilisation de l'argumentation comme cadre unifié sous-jacent pour (1), (2) et (3) pourrait alors ouvrir la voie à une intégration harmonieuse des différentes capacités juridiques au sein de nos « bots juridiques » unifiés, reflétant la façon dont les juges combinent les lois et les cas antérieurs/jurisprudences pour prendre leurs décisions. Cela conduira à fournir également des suggestions transparentes et vérifiables aux utilisateurs (les juges) renforçant ainsi la confiance dans ces systèmes.

Les « bots juridiques » pourraient alors apparaître comme des compagnons, plutôt que comme des concurrents pour les décideurs juridiques. Ces derniers seront les utilisateurs actifs de ces « bots juridiques » et pourraient ainsi à terme interagir avec eux en fournissant notamment des commentaires (un feedback) sur leur comportement et leurs résultats. Par exemple, ils pourront spécifier si une règle ou un cas antérieur/jurisprudence est correctement extrait du texte (et donc fournir un feedback à (3)), si une recommandation basée sur le droit (les lois) est correcte (et donc fournir un feedback à (1)) ou si un cas antérieur/jurisprudence est en effet pertinent pour le cas présent (et donc fournir un feedback à (2)). Cela permettrait alors de réinjecter de l'information dans le système et d'améliorer ses performances dans le temps.

Les "bots juridiques » offrent à ce titre plusieurs nouveautés par rapport aux systèmes antérieurs présentés dans la littérature (voir par exemple ASHSD-II, CHIRON, JUDGE, Split-Up, SHYSTER, TAXMAN). Ces nouveautés sont: 1) l'intégration agile de différents types de raisonnement en utilisant l'argumentation comme moyen commun, basée sur la semi-automatisation de l'extraction d'arguments à partir de textes juridiques et la génération automatisée d'arguments fondés sur des règles/préférences à partir des cas antérieurs et 2) l'amélioration continue du système par le biais de l'interaction avec l'utilisateur (par exemple, les avocats et les juges), le feedback facilité par le haut niveau déclaratif et naturel de l'argumentation, tout à fait dans l'esprit juridique traditionnel. Ainsi, par exemple, les juges ne seront pas devant des recommandations du type « boîte noire » ou des statistiques brutes avancées comme preuves pour les recommandations proposées, mais des justifications articulées à leur place, basées sur l'analyse des données et les interactions avec les utilisateurs juridiques.

3) L'IA peut-elle faciliter le processus de médiation et plus particulièrement le règlement des conflits (en automatisant le processus)?

Des développements importants concernant des techniques d'interaction et de dialogues automatisés entre plusieurs entités computationnelles (appelées « agents intelligents ») fondés sur des techniques systèmes multi-agents, devraient permettre dans un futur proche de pouvoir proposer des solutions concernant l'automatisation (ou semi-automatisation) du processus de médiation.

En particulier, grâce à l'utilisation sous-jacente des méthodologies d'argumentation, ces bots pourraient être naturellement déployés pour le support automatisé de la médiation des différends entre les parties adverses. En présence de conflits (par exemple, concernant des accidents de la circulation, des accidents médicaux, des divorces, des conflits du travail, etc.), des « bots juridiques » pour des différentes parties, instanciés avec l'expertise associée au domaine du conflit, échangeront des faits et construiront des arguments à partir de leurs propres points de vues. Ensuite, les bots impliqués au conflit pourront débattre ouvertement en échangeant des arguments basés sur les informations fournies, éventuellement en collaboration avec leurs utilisateurs humains. Dans de tels débats, un tiers neutre (le médiateur, par exemple un juge) pourrait obtenir la trace du débat et décider pour le clore s'il convient d'accepter les déclarations faites par les opposants au cours des échanges. Les parties adverses pourraient alors profiter de ce système pour avoir des estimations quant à la possibilité que les prémisses de leurs arguments soient acceptées dans un procès réel. Cela pourrait conduire soit à un règlement à l'amiable sans un vrai procès entre les parties opposées, soit à un procès réel devant les tribunaux. Dans le second cas, comme les "bots juridiques » pourront fournir une trace de leur raisonnement (les lois et les cas antérieurs/jurisprudences) qu'ils ont utilisés dans leur débat, les avocats humains des parties adverses, ainsi que les parties elles-mêmes, pourraient tirer profit de ces informations pour préparer leur défense dans le procès devant les tribunaux.

4) Comment l'IA peut-elle transformer le fonctionnement de la justice

Le développement de cette IA présente des bénéfices sociétaux indéniables. Elle permettrait de diminuer le nombre de recours et donc de lutter contre le phénomène de judiciarisation constaté. À ce titre, elle offrirait indéniablement une réponse au désengorgement des juridictions judiciaires et administratives. En rendant plus facile l'analyse de l'information et en croisant les données, elle offrirait un gain de temps et une diminution des coûts importante, ce qui faciliterait l'accès au droit et à la justice et contribuerait à les démocratiser. Elle aurait le mérite de renforcer la sécurité juridique en donnant la possibilité aux justiciables, à l'avocat, mais aussi au juge, de pouvoir anticiper toute solution du différend. Elle permettrait également de manière non plus ex-ante, mais en quelque sorte ex-post, d'assurer un contrôle « qualité » des décisions rendues par les juges. Ainsi, elle pourrait conduire à la diminution aussi bien du nombre d'appels interjetés que du nombre de recours en cassation exercés devant les juridictions concernées. Elle présenterait également l'avantage de rendre effective les dispositions de la loi de modernisation de la justice du XXIe siècle du 18 novembre 2016 qui visent à favoriser non seulement l'accès à la justice et les modes alternatifs de règlement des différends, mais aussi à recentrer les juridictions sur leurs missions essentielles. Ainsi, elle permettrait de répondre à l'exigence de ce texte qui dans un certain nombre d'hypothèses rend la médiation obligatoire préalablement à toute saisine judiciaire.

Certes, un tel développement ne sera pas sans soulever un certain nombre de questions importantes par rapport notamment à l'impact qu'il pourrait avoir non seulement sur la profession d'avocat, sur la conception du rôle du juge mais aussi sur la nature même des décisions de justice. Plus largement, il contribuera à transformer le fonctionnement même de la justice. Il conduira ainsi à repenser la formation d'un grand nombre de professionnels du droit, voire également leurs règles professionnelles ou déontologiques.

Les techniques développées seront aisément mise en œuvre dans les domaines où la part de subjectivité est limitée (application des règles de prescription, exigence de formalisme, montant des indemnités versées dans divers contentieux…) mais elles sont également susceptibles d'intégrées une part de subjectivité dans l'analyse.

L'intérêt stratégique pour la France de développer ce genre d'outils est non seulement économique au regard notamment du développement considérable de la part des legaltechs dans le marché du droit (cf. Etats-Unis), mais aussi pour renforcer la position du droit français dans un monde globalisé. En effet, une IA construite sur la base du droit français permettrait de renforcer la position du droit continental face au développement de la common law dont la prégnance, voire la suprématie dans le domaine du digital est aujourd'hui indéniable.

Si pour l'instant aucune legaltech ne semble s'être orientée vers le développement de ce type d'IA, il est important de soutenir les initiatives qui en France permettraient de parvenir à celui-ci. En effet, la réalisation de ce type d'IA est d'autant plus envisageable que le droit français présente l'avantage de permettre un accès libre des données, que ce soit des textes juridiques mais aussi, d'un nombre important de décisions de justice.

Comme cela a été souligné (voir points (2) and (3)) les solutions techniques que les recherches en matière d'IA proposent sont pour certaines très matures et permettent d'envisager précisément et avec certitude le développement à très court terme de recherches appliquées permettant de faire apparaître les « bots juridiques » évoqués. Ces "bots juridiques" seront un soutien important pour les utilisateurs juridiques (par exemple les avocats et les juges) car ils pourraient leur permettre de vérifier la solidité de leurs décisions en les confrontant avec des recommandations semi-automatisées à partir des raisonnements fondés sur des cas antérieurs (jurisprudences), des textes et des lois. De plus les "bots juridiques" seront bénéfiques à tous les utilisateurs dans le domaine du droit, y compris les non-experts juridiques. En particulier, ces bots pourraient être naturellement déployés pour le support automatisé de la médiation des différends entre parties adverses.

Néanmoins, la réalisation de ce projet n'a pour l'instant pas été possible car elle se heurte à différents obstacles. Le premier obstacle est celui du développement de recherches pluridisciplinaires permettant de croiser des compétences en droit avec des compétences en intelligence artificielle. Or par des projets de recherche antérieurs, les porteurs du projet ont déjà su réaliser ce difficile rapprochement entre juristes et informaticiens spécialisés dans l'IA. La réalisation du projet suppose également que puisse être fléchés, au niveau national, des fonds permettant à des équipes pluridisciplinaires de pouvoir développer ce type de recherches alors que pour l'heure, ni le programme européen (H2020), ni le programme ANR pour 2018 ne prévoient de programmes blancs ou fléchés permettant le financement de ce type de recherches. Or, pour autant le financement de ces recherches visant à la création à terme de ce type d'IA Droit est essentiel alors même qu'aucune legaltech ne s'est pour l'instant engagée dans ce domaine. Compte-tenu des enjeux stratégiques soulignés, il est important que la France puisse être à la pointe de l'innovation dans la réalisation de ce type d'IA.

Recommandation

Au regard des enjeux stratégiques pour la France, tant d'ordre économique que de nature à assurer l'influence du droit continental par rapport à la *common law*, et plus particulièrement du droit français, il est essentiel à ce stade du développement des recherches en matière d'IA en France d'allouer des financements à des équipes de recherches pluridisciplinaires prêtes à réaliser des recherches appliquées en matière d'IA et droit. Au-delà des opérations de justice prédictive déjà proposée, l'IA pourrait offrir une modélisation fidèle du raisonnement juridique. Cela permettrait un mode de résolution automatisée des litiges en faisant intervenir l'IA comme interprète de l'opposition des arguments des parties, une aide à la décision des juges et une automatisation possible des débats juridiques entre intelligences artificielles représentant des parties adverses et assistant leurs avocats. Ce type d'IA serait une réponse à la judiciarisation de la société, elle permettrait de renforcer la sécurité juridique, de favoriser l'accès à la justice ainsi que le règlement amiable des litiges.

Membres du groupe

BERLIOZ Pierre, Professeur, Faculté de Droit, Université Paris Descartes, CEDAG
BONZON Elise, MCF, LIPADE, Université Paris Descartes

GENTY Antoine, Avocat au Barreau de Paris
KAKAS Antonis, Professor, Dept. of Computer Science, University of Cyprus (Cyprus)
LAUDE Anne, Professeur, Faculté de Droit, Université Paris Descartes
MORAITIS Pavlos, Professeur, LIPADE, Université Paris Descartes
MORLET Lydia, MCF, Faculté de Droit, Université Paris Descartes
TABUTEAU Didier, Professeur Associé, Faculté de Droit, Université Paris Descartes, Conseiller d'Etat
TONI Francesca, Professor, Dept. of Computing, Imperial College (UK)

GROUPE DE TRAVAIL 1.2
FORMATION

Groupe Formation

Table des matières

1. Périmètre d'étude du groupe de travail

Préliminaires : Définition et écosystème thématique

Définition. Nous envisageons l'IA comme un usage visant à automatiser les tâches et passant par l'intégration de plusieurs briques technologiques relevant d'un écosystème thématique.

Nous définissions l'*écosystème thématique* de l'IA ainsi :

- **Information** : l'organisation et la gestion numérisées de la donnée ou de l'expertise (procédures de capture, bases de données).
- **Matériel** : l'infrastructure matérielle comprend :
  - les capteurs
  - la technologie de calcul (HPC et embarqué)
- **Algorithmes** : les algorithmes proprement dits, les stratégies de résolution, les stratégies de calcul, les systèmes experts, etc.
- **Interaction** : la robotique, les interfaces homme-machine (IHM), l'ergonomie
- **Encadrement : les savoir-faire transverses permettant l'implémentation contrôlée de ces technologies i.e. droit, philosophie et éthique, acceptabilité et mesure de l'impact individuel et social et plus largement les sciences humaines et sociales (SHS**

Mission. Définir une stratégie nationale en intelligence artificielle afin de positionner la France comme un acteur majeur sur le domaine passe par la définition des besoins en formation sur ce thème et leur priorisation.

Démarche. Notre travail s'est organisé autour de deux axes :

- développer la formation de spécialistes en IA au sens de créer des experts dans les différents thèmes définissant l'IA (cf. écosystème thématique)
- préparer tous les citoyens à saisir les opportunités de cette nouvelle révolution et à interagir avec des systèmes dotés d'IA, tant comme utilisateur que comme expert "assisté" ou "augmenté" par une IA.

Méthode/Sources/Groupe de référence. Enfin, nous avons souhaité aborder ces besoins de formation pour les acteurs privés mais aussi pour les acteurs publics en nous appuyant sur une cinquantaine d'entretiens dont 40% de startups.[1]Nous avons également diffusé un questionnaire à plus de cent cinquante entreprises.[2][3]

2. Etat des lieux

2.1. Maturité des acteurs et expression des besoins inhomogènes sur le sujet IA

Startups. Les startups sont les premiers acteurs de l'IA en France. Plus de 200 d'entre elles intègrent ou conçoivent des systèmes d'IA[4]. Elles sont conscientes des promesses de ces technologies et les premiers consommateurs d'expertise dans ce domaine. Ce sont les premiers acteurs de réorientation professionnelle en proposant à des professionnels de devenir des entraîneurs ou superviseurs d'IA.

GE/ETI. Les grandes entreprises accélèrent l'automatisation de leur processus interne et expérimentent les possibilités offertes par les dernières technologies en suivant les "modes" sans vraie connaissance

[1] listées en annexe 6.1
[2] listées en annexe 6.2
[3] Pour l'analyse des résultats voir l'annexe 2
[4] Recensement effectué en 2016 par P. Strachman, (ISAI) pour l'évènement "France is AI"

de l'impact ou valeur pour leur société. Selon leur désir de maîtriser ces technologies, elles se tournent vers les solutions clés-en-main des géants de l'informatique (IBM, Microsoft, etc) ou vers les briques technologiques portées par des startups.
La plupart des entreprises sondées, y compris celles qui se déclarent peu matures sur le sujet, intègrent quelques briques d'IA dans leur processus métier et/ou leurs systèmes de relation client, en particulier dans les domaines Banque / Assurance. Les seules à avoir des équipes spécialisées sont plutôt des entreprises IT.

PME. Les PME et TPE (hors startups) semblent encore loin de ces sujets et en pleine transformation numérique.[5]

Administrations. Les administrations publiques poursuivent leur transformation par l'action du SGMAP, notamment à travers la DINSIC. Le besoin d'automatisation semble être une réponse aux attentes de la LOLF mais le sujet reste peu abordé.

Elus. Les politiques et leurs administrés sont peu préparés pour appréhender les enjeux du sujet.

2.2. Les besoins de compétences autour de l'IA font apparaître différents besoins de formation et de sensibilisation des décideurs

Notre étude part du postulat que l'automatisation généralisée rendue possible par l'IA dans tous les domaines de l'activité humaine aura un impact massif sur les organisations, entreprises ou administrations, et sur certaines fonctions au sein de celles-ci : du décideur aux utilisateurs d'IA, y compris les ingénieurs R&D. [6]

D'abord, les enjeux liés à la compétitivité des entreprises et la prévention du risque d'intermédiation de certains acteurs traditionnels vont créer une forte demande en concepteurs de systèmes d'IA.

Décideurs. Afin que l'entreprise, l'administration et les politiques se saisissent de ce nouvel enjeu et décident des moyens à y consacrer, **la sensibilisation des décideurs** (et managers) [7] est une étape à court terme importante pour identifier les enjeux stratégiques et permettre aux organisations de s'adapter aux transformations auxquelles elles se trouveront exposées.

Utilisateurs et chefs de projet. Il y a des besoins de compétences particulières pour des personnes déjà présentes dans les entreprises ou des personnes souhaitant se réorienter professionnellement. Ils vont concerner essentiellement deux métiers :

- des développeurs informatiques appelés à être des intégrateurs d'IA
- des spécialistes « métier » capables d'organiser le problème à résoudre pour le rendre accessible aux solutions techniques existantes, sélectionner et utiliser les nouveaux outils embarquant une IA et parfois d'entraîner et de superviser une IA.

On attend également une forte demande de chefs de projets intervenant de manière transversale sur le développement, l'intégration, et la maintenance des systèmes d'IA.

---

[5] https://www.lesechos.fr/pme-regions/actualite-pme/0211781271292-numerique-dans-les-pme-les-acteurs-en-appellent-a-un-plan-durgence-2064360.php

[6] On trouve une analyse similaire dans le rapport « Artificial Intelligence, Automation, Economy » remise au président des Etats-Unis en 2016.

[7] Dans un sondage sur 1600 entreprises de plus de 1000 employés, plus de 50% des sondés citent le « manque de compréhension des bénéfices que l'IA peut apporter » et « le manque de compétence en interne pour mettre en place et manager ces projets » comme les barrières principales à un plus grand impact de l'IA (« AI Maturity Index» Infosys 2016).

Concepteurs. Les **formations initiales d'expertise existantes semblent satisfaire** les entreprises en France et à l'international en terme de qualité et niveau d'expertise. Cependant, la population des ingénieurs spécialisés en IA suffit à peine à combler les attentes de startups et de quelques grands groupes présentant les problèmes les plus attractifs. On observe notamment un mouvement conséquent de départs à l'étranger (notamment vers les Etats-Unis) des jeunes ingénieurs et chercheurs, ce qui soulève la question de l'attractivité des écosystèmes français.

2.3. Formations existantes en réponse aux besoins

2.3.1. Une formation initiale offrant un bon niveau d'expertise pour les ingénieurs en IA mais encore morcelée thématiquement et géographiquement

Les formations de pointe identifiées sont principalement au niveau master. Les critères définissant une formation attractive en IA et répondant aux besoins des entreprises relèvent principalement de l'accès à des profils généralistes de haut niveau, et de la rencontre de quatre différents facteurs :

Environnement thématique. Toutes les composantes techniques (cf. section 1.) de l'IA sont représentées et de qualité mais le lien avec la chaîne de valeur de l'IA n'est pas encore suffisamment mis en avant et souffre d'une approche trop verticale, là où les professeurs des universités américaines présentent une double affiliation[8] avec un rattachement principale dans le cœur d'expertise et un rattachement secondaire marquant un domaine d'application.

Bassins de recrutement. Les profils les plus adaptés pour irriguer ces formations initiales sont principalement ceux des ingénieurs généralistes. La France bénéficie grâce à son dispositif de Grandes Écoles d'un contingent significatif de tels profils mais qui se trouvent concentrées en région parisienne.

Écosystème industriel. Les meilleures formations bénéficient d'un écosystème industriel varié et en attente d'applications des dernières technologies pour la résolution de leur problème. Il offre à la fois des débouchés attrayants aux meilleurs élèves et stimule les enseignants chercheurs dans l'application de leurs recherches. Les startups présentent la meilleure attractivité pour les étudiants mais ne sont pas suffisamment présentes sur l'ensemble du territoire.

Écosystème de recherche. La recherche française en mathématiques se positionne au plus haut niveau à l'international, notamment dans le domaine du *machine learning*. Les formations au niveau master recherche bénéficient potentiellement d'intervenants de haut niveau, de contenus en phase avec l'actualité scientifique, les innovations technologiques et d'exemples illustrant ces applications.[9]

2.3.2. Une offre de formation naissante dans le domaine des SHS attendue sous forme de module dans les cursus d'ingénieur IA mais prématurée pour les entreprises

Les formations en SHS autour de nouveaux enjeux posés par le développement de l'IA sont encore rares. Si d'un point de vue industriel, une production massive d'expertise sur ces composantes semble un peu prématurée (sauf pour les GAFAM[10]), certaines formations d'ingénieurs et de décideurs cherchent des intervenants pour des modules de sensibilisation, notamment droit, philosophie, design, etc.

---

[8] C'est le modèle du "joint appointment". Par exemple ici la description de celui du CMU : http://www.cmu.edu/mcs/fac_staff/handbook/joint_courtesy.html

[9] La France se place au 3e rang mondial en nombre de publications en mathématiques les plus citées à 2 ans et au 1er en nombre de publications dans le cadre de collaborations internationales (Source : "L'impact socio-économique des mathématiques en France" p11-12 AMIES 2015)

[10] Acronyme désignant les "Géants du Web", Google, Apple, Facebook, Amazon et Microsoft qui sont les cinq grandes firmes américaines qui dominent le marché du numérique

2.3.3. Des outils de formation continue peu fournis dans le domaine de l'IA et des difficultés à standardiser des contenus pédagogiques

Les principaux outils identifiés pour la formation continue sont les suivants : (i) formation interne spécialisée, (ii) formation à distance (type MOOC), (iii) formation diplômante présentielle (type CNAM, mastère professionnel).

Modules IA de référence. Les formations portées par des plateformes américaines telles que Coursera et Udacity contribuent actuellement à former une partie des intégrateurs d'IA et ainsi à répondre au manque d'experts disponibles sur le marché. Des initiatives nationales commencent à émerger essentiellement sur le sujet Big Data (OpenClassroom, FUN).

Une offre nationale encore peu fournie. Des offres de formations internes et diplômantes ont été développées avec succès sur des domaines connexes à l'IA (par exemple, le Master spécialisé Big Data Telecom ParisTech) mais on peine à identifier des modules spécifiques sur l'IA et l'on peut craindre une floraison de formations en IA avec des qualités très variables. La durée de certification des formations auprès du CPF peut être vu comme un frein au déploiement de nouvelles formations et donc à la formation des entreprises (ex : un sondage d'insertion professionnelle auprès de 3 promotions doit être mené ce qui pour les formations longues peut repousser l'ouverture de ces formations auprès des organismes de formation professionnelle).

Une difficulté à standardiser des contenus pédagogiques. L'environnement technologique très mouvant et très riche est une réalité en décalage avec la temporalité du système de formation continue en entreprise (par exemple, les certifications enregistrées au CPF voient leur référentiel de compétences figées pour plusieurs années - cf. rapport FAFIEC).

2.4. Des moyens de formation à développer en profondeur et une attractivité nationale à construire sur la discipline

Accélérer le développement de l'IA en France et son appropriation par tous les enseignants et chercheurs désireux de s'en donner les moyens, c'est offrir les ressources pédagogiques leur permettant de se former et d'expérimenter rapidement les technologies de l'IA.

Les données jouent un rôle central pour le développement de l'IA. Là où les Etats-Unis comptent des **bases de données**, gratuites et libres, en langue anglaise, les enseignants/chercheurs ne disposent que de quelques jeux de données, anciens, et non mutualisés en français.

Accès aux infrastructures. Les méthodes d'apprentissage statistique nécessitent des **infrastructures de calcul particulières** (GPU) qui peuvent être rapidement coûteuses pour un formateur dont l'établissement n'est pas équipé.

Kits pédagogiques. Si **les tutoriels et les librairies de calcul** accessibles par internet se multiplient, les supports développés en France restent rares. Des initiatives de renommée internationale comme Scikit-Learn[11] ou IPOL[12] bénéficient de trop peu de soutien pour se développer et développer des formations pour les enseignants. On peut étendre cette idée aux kits 'robotique'.

Recrutement des formateurs. Plusieurs difficultés apparaissent : (i) défaut de compétences sur les

[11] C'est la librairie de référence pour l'apprentissage statistique dans le langage Python avec 200,000+ ainsi que la plus utilisée pour l'enseignement dans nos grandes écoles et nos universités (http://scikit-learn.org/).
[12] Image Processing On Line : Portail de référence pour la recherche reproductible en traitement d'images avec 200,000+ expériences réalisées en ligne (http://www.ipol.im/).

interfaces IA-métier (exemple, droit et IA), (ii) valorisation de l'investissement pour construire des contenus pédagogiques, (iii) fuite des cerveaux à très haut niveau résultant des différentiels de salaires (public-privé et par rapport à l'étranger), de la déstabilisation des institutions liées au regroupement des établissements d'enseignement supérieur et de recherche et d'un contexte administratif inadapté à la mise en place d'équipes durables et de moyens techniques ambitieux.

2.5. Leviers d'acceptabilité et d'intégration de l'IA dans l'économie

2.5.1. L'IA pour tous : premier et second degré

Un pays préparé à intégrer pleinement cette nouvelle technologie dans son économie est un pays dont les citoyens comprennent les enjeux liés aux données et ont la capacité de comprendre, lire, utiliser, interpréter et communiquer des données afin de participer au débat public de manière éclairée. Former des experts de l'IA passe donc par la sensibilisation et la formation de tous les élèves et étudiants français à la science informatique de l'école primaire à l'enseignement supérieur. Au lycée, la spécialité « Informatique et science du numérique » (ISN) trouve sa place progressivement. Les intervenants sont principalement des enseignants de mathématiques qui reçoivent une formation spécifique.

L'enseignement de la programmation arrive au collège. La réforme des collèges mise en place en 2016 donne une place nouvelle au codage et aux bases de la programmation informatique dans les programmes d'enseignements en mathématiques. La programmation intervient au moyen d'interfaces de programmation visuelle et ludique, comme celles proposées par Scratch, adaptées à des enfants du collège.

L'apprentissage de l'anglais. L'anglais reste la langue première de l'informatique et l'introduction d'une première langue vivante au CP dans le cadre de la réforme est une évolution très positive.

La pratique de la robotique est abordée dans les programmes du premier degré. Depuis 2015, les nouveaux programmes des enseignements du premier degré intègrent largement le numérique, y compris des rudiments de programmation et de robotique.

Pourquoi ne pas aller plus loin et plus vite ? Il est encore tôt pour faire un état des retours d'expérience sur l'implémentation de ces réformes mais un certain nombre de propositions sont déjà formulées pour les étendre : identification d'une discipline spécifique dans le primaire, formation des formateurs à renforcer, prise en compte des aspects liés aux données, initiation technologique au support physique (structure d'un ordinateur).

2.5.2. L'IA au service de l'enseignement

Florence Robine, Directrice générale de l'enseignement scolaire, a déclaré récemment : "les neurosciences, l'intelligence artificielle et le "big data" vont, dans les années qui viennent, bousculer notre compréhension de la manière dont les élèves apprennent et améliorer l'évaluation". Ainsi, l'IA sera bien plus qu'un nouveau contenu à enseigner dans la mesure où il pourrait être utilisé un jour pour personnaliser le programme de formation d'élèves. Pour se préparer à cette transformation des manières d'enseigner « l'usage du numérique doit être systématisé dans les formations et l'évaluation des étudiants»[13]

L'IA au service de toutes les disciplines. On note de nombreux cas d'utilisation de l'IA dans l'apprentissage des langues par exemple, avec des applications telles que *Duolingo* ou *Iknow*[14] qui utilisent notamment les méthodes de mémorisation dites de répétition espacée dans l'apprentissage du

[13] Rapport « Pour une société apprenante, proposition pour une stratégie nationale de l'enseignement supérieur du XXIe siècle », (STRANES 2015).
[14] http://iknow.jp/

vocabulaire. La possibilité de créer des exercices à l'infini et en fonction de la progression des élèves semble décomplexer l'élève dans son apprentissage.

L'IA pour personnaliser le parcours d'enseignement. Dans certaines écoles d'informatique privées (42, Epitech), des IA sont utilisées pour évaluer le rendu des élèves de manière objective tout en personnalisant les pistes d'amélioration.

2.5.3. Le rôle de l'Etat

L'Etat par ses administrations a un rôle clé dans l'intégration et l'acceptabilité de l'IA dans l'économie et doit faire preuve d'exemplarité. Comme le note le Conseil National du numérique : «le développement de nouveaux usages numériques dynamite le fonctionnement traditionnel de l'administration et la contraint à impulser en interne une dynamique de changement »[15].

L'Etat peut fixer un cap. L'appropriation de cette technologie pour le compte des institutions et du fonctionnement de l'administration est un message positif à envoyer aux administrés comme il l'a été pour d'autres technologies (internet, le prélèvement bancaire, la signature électronique, etc.).

La préparation des cadres de la fonction publique. L'IA peut être une source d'opportunités et de risques pour les administrations ; les fonctionnaires doivent donc être en capacité de s'en saisir.
Les formations données aux fonctionnaires stagiaires dans les principales écoles de la fonction publique abordent depuis 1-2 ans la place des données dans l'évaluation des politiques publiques et présentent le numérique comme un levier possible d'amélioration des services de l'Etat (et comme un élément du quotidien de ces fonctionnaires). Toutefois, le sujet de l'IA n'y est pas encore abordé et pas plus dans la formation continue.

Certains corps techniques (INSEE et MINES) forment leurs fonctionnaires stagiaires aux enjeux des big data mais n'ont pas développé d'expertises particulières sur le sujet IA.

4. Enjeux et recommandations

4.1 Sensibiliser le grand public et les décideurs aux enjeux de l'IA

Un pays préparé à intégrer pleinement cette nouvelle technologie dans son économie est un pays dans lequel ses citoyens comprennent les enjeux liés aux données et ont la capacité de lire, utiliser, interpréter et communiquer des données afin d'en tirer parti pour leur activité et de participer au débat public de manière éclairée.

De plus, afin que les entreprises, l'administration et les politiques se saisissent de cette nouvelle révolution et décident des moyens à y consacrer, la sensibilisation des décideurs (et managers) est une étape importante à court terme, pour permettre aux organisations de s'adapter pro-activement aux transformations inévitables de leurs méthodes de production et de travail.

<u>Recommandation : Créer un « plan IA» piloté par l'Agence du Numérique</u>

Un plan national de sensibilisation, développé sur au moins 2 ans, permettrait de définir des actions de communication et de formations de tous publics en commençant prioritairement par les décideurs publics et privés. Ses missions pourront être de :

- Créer des contenus pédagogiques pour les dirigeants privés ou publics sous la forme de mini-

[15] Rapport «Ambition numérique : pour une politique française et européenne de la transition numérique » p99 Conseil national du numérique (2015)

MOOC sur l'IA pouvant être associés à des ateliers organisés dans des tech-places
- Intervenir auprès des organismes représentatifs (MEDEC, SYNTEC, etc.)
- Mettre en place une campagne nationale de sensibilisation (tous médias) à destination du grand public
- Mettre en place un site d'information officiel, destiné au grand public, avec des démonstrations et des explications de vulgarisation sur le sujet

Recommandation : Introduction de modules techniques dans les écoles de management et des cadres de l'Etat

Pour former des décideurs consciences de ces enjeux, il faut inciter ces acteurs à :

- mettre en place de modules d'initiation IA pour leurs étudiants (actuellement focalisés sur les thèmes data analytics, big data, etc)
- développer des programmes à destination des futurs créateurs de startups IA ou futurs chefs d'entreprise acheteurs de solutions d'IA. [16]

Exemple de fiches de formation en annexes 5.1, 5.2, 5.3.

4.2 Préparer les citoyens de demain à concevoir et interagir avec des IA

Dans une vision prospective autour de l'IA, on peut s'attendre à ce qu'elle occupe place centrale dans les sociétés et les organisations dans un futur proche et les jeunes générations doivent y être confrontées aussi tôt que possible à travers l'enseignement. Ceci passe notamment par l'enseignement de l'informatique au plus grand nombre afin de créer la culture nécessaire à chaque citoyen pour utiliser, comprendre et concevoir les futures IA.

Recommandation : Instauration d'un enseignement "IA, traitement des données et sciences numériques" de l'école primaire au lycée

Cet enseignement, critique, multiforme et pluridisciplinaire, doit montrer que les aspects techniques sont interdépendants des enjeux éthiques, sociétaux et des applications. Cet effort sera naturellement amené à se diffuser dans toute la population. Il s'adosse aux apprentissages en mathématiques et implique l'utilisation d'applications et une initiation à la programmation. Cinq moyens d'actions sont identifiés :

- Un enseignement « IA, traitement des données et sciences numériques» au primaire ;
- Création de postes «référents numériques» et affection sur compétences spécifiques en collèges et lycées ;
- Création et valorisation de contenus pédagogiques ;
- Implication des enseignants de toutes disciplines pour un usage des outils de l'IA ;
- Développement d'enseignements spécialisés dans le second degré vers les métiers du numérique

Cette recommandation est détaillée dans l'annexe 4.

4.3 Développer l'écosystème des formations en IA en France afin qu'elles soient attractives et puissent se développer

Le succès des formations en IA dépend directement des interactions qu'elles pourront développer avec

[16] Columbia a déjà pris une initiative de ce type et formé 2000 personnes avec une majeure IA (dont 300 français)

les industriels, porteurs de problématiques technologiques et économiques, et les laboratoires de recherche qui détiennent un savoir-faire et des outils pour les résoudre.
D'autre part, les usages au quotidien de l'IA ouvrent de nouveaux champs de recherche pour les sciences humaines et sociales, le droit et la philosophie où la France a la capacité de se positionner en précurseur sur certaines spécialités. Elles occuperont une place significative dans la conception et l'intégration maîtrisée des IA. Le développement de ces expertises pourrait rapidement contribuer à créer un avantage compétitif national, comme l'est par exemple la législation pionnière sur les drones.

Recommandation : Inciter les interactions entre les centres de recherche, centres de formation et entreprises autour de l'IA pour favoriser l'innovation et l'attractivité de l'écosystème

- Stimuler le lien entre les entreprises et les centres de formation en incitant à la participation d'ingénieurs spécialistes en IA dans des cours ou des formations.
    - Le développement du mécénat de compétences permettrait à des entreprises de s'impliquer plus en profondeur dans les formations.
- Pour répondre au problème de la « fuite des cerveaux » extrêmement présent dans ces disciplines, nous proposons de renforcer la possibilité pour les chercheurs de travailler dans l'industrie de manière saisonnière. Bien que des dispositifs existent, tels que la mise à disposition à temps partiel, ils sont encore peu utilisés car mal reconnus par les pairs et les administrations des établissements.
- Nous recommandons de développer prioritairement des formations en IA dans des parcours par filière industrielle en particulier, dans des campus au centre de bassins d'emplois et de pôles d'excellence de recherche. Nous proposons notamment de favoriser le lancement de formation en IA au sein des campus des métiers et des qualifications notamment :
    - des formations de sensibilisation pour les responsables de ces campus et les industriels du réseau (1 journée) ;
    - des ressources afin de développer ces formations :
- de modules de formation non présentielle (ex: Licence OpenClassRoom) à privilégier pour pallier à l'absence de moyens pour accueillir des enseignants chercheurs sur ces spécialités ;
- des accès à des moyens de calcul (GPU) et à des corpus d'apprentissage nationaux pour les enseignants, chercheurs et étudiants mais pour les industriels du campus afin de les aider à monter en compétence et développer des débouchés ;
    - Les modules de formation en IA peuvent être proposés dès la licence L3 développée sur le campus et un parcours ingénieur IA devra être fléché au sein de la filière.

Recommandation : Encourager l'émergence de travaux de référence en droit, éthique, SHS relatifs aux questions nouvelles posées par l'IA

Pour former une génération de chercheurs pionniers sur ces domaines, il est nécessaire de créer, à court-terme, des débouchés pour ces experts IA-SHS, en dehors des seules « GAFAM ». Ce soutien peut prendre la forme de :

- soutien financier au travers de bourses de thèses spécifiques fléchées en IA
- création d'une chaire ou d'un observatoire qui regrouperait des chercheurs ou des projets de recherche travaillant sur des questions posées par l'impact de l'IA en SHS
- aides spécifiques à la publication ou à la diffusion de travaux en SHS relatifs à l'IA

- mise en place dans les ANR d'un volet spécifique sur les travaux en SHS relatifs à l'IA
- mise en place dans les PIA technologiques relatifs à l'IA d'un volet spécifique sur les impacts sociétaux.

4.4 Développer les moyens techniques pour la formation en IA

Développer les formations en IA nécessite de mettre à disposition des enseignants et des étudiants les ressources incontournables du développement de nouveaux systèmes. De plus, ces ressources peuvent aisément être mutualisées entre les acteurs (recherche, formation, entreprises). Là où les Etats-Unis comptent des **bases de données**, gratuites et libres, en langue anglaise, les enseignants/chercheurs ne disposent que de quelques jeux de données, anciens, et non mutualisés en français. De plus, **les tutoriels et les librairies de calcul** accessibles par internet se multiplient et les supports développés en France bénéficient de trop peu de soutiens pour se développer et proposer des formations pour les enseignants (Scikit-Learn est un exemple notoire).

<u>Recommandation : Développer une ressource nationale (voire de la francophonie) de données non agrégées, structurées et labellisées en Français</u>

- Créer une base de données structurées, non agrégées, labellisées, de toutes natures (voix, images, textes, signaux, vidéos, etc.) pouvant s'appuyer sur les ressources publiques non ou faiblement exploitées (ex : Base INA, Base enregistrement RadioFrance, Sous titrage de la télévision publique, Santé) ;
- Placer le développement et la gestion de cette base de données sous l'égide de data.gouv.fr dans la continuité des travaux d'Etalab ;
- Définir des mécanismes durables pour l'alimenter et structurant des communautés et des méthodes pérennes de labellisation autour de ces nouvelles ressources ;
- La structuration et la mise à disposition de ces données pourront faire l'objet d'appels au PIA 3 ;
- Elle doit être ouverte gratuitement aux JEI afin d'inciter le développement de startups en IA en France ;
- Allonger la durée de détention des données personnelles imposée par la CNIL.

<u>Recommandation : Pérenniser les initiatives logicielles open source française en IA bénéficiant à la recherche et à la formation</u>

En effet, nos initiatives existantes, même les plus reconnues, ont du mal à obtenir des financements dans le cadre des AAP. Il est donc nécessaire de :

- créer via le PIA, un fond dédié spécifiquement au financement des logiciels open source à intérêt public pour :
    - Consolider les logiciels ayant du succès en finançant le développement des plateformes sur le modèle de ce qui existe aux Etats-Unis (ex : CodeOcean), où les logiciels ayant fait leurs preuves ne sont plus vus comme des livrables de fin de projet, mais des actifs valorisés et entretenus ;
    - Favoriser leur dissémination en finançant des évènements de formation (ateliers, workshops) à travers le réseau des écoles et des universités. Pour faciliter l'organisation de tels workshops le Center for Data Science (CDS) de l'université Paris-Saclay a financé le développement d'une plateforme de travail collaboratif

appelée RAMP[17] ;
    - Renforcer le rôle des SATT dans l'accompagnement des projets open source afin de les aider à trouver un modèle économique pérennisant leur activité
- Favoriser l'achat de licences (ou de dons) pour les logiciels opensource par les entreprises.

4.5 Faire de l'Etat un pionnier dans l'adoption de l'IA

L'Etat par ses administrations a un rôle clé dans l'intégration et l'acceptabilité de l'IA dans l'économie. L'appropriation de cette technologie pour améliorer le fonctionnement de l'administration et des institutions est un message positif à envoyer aux administrés comme il l'a été pour d'autres technologies (internet, le prélèvement bancaire, etc.). Les gains économiques et en qualité de service public envisageables sont importants mais ils nécessitent de sensibiliser largement les fonctionnaires à ces enjeux.

Dans l'éducation, l'IA est un outil particulièrement prometteur qui peut favoriser la personnalisation du parcours de formation (pour l'accompagnement et l'évaluation), l'amélioration de la mémorisation des savoirs (méthode de répétition espacée) et la création d'exercices d'entraînement pour toutes les disciplines. Cette utilisation pédagogique transmet de plus une vision positive de l'IA qui peut participer fortement à son acceptation et à sa diffusion dans la société

<u>Recommandation : Création d'un pôle d'innovation EdTech au sein du Ministère de l'Education Nationale et de l'Enseignement Supérieur et la Recherche</u>

Ce pôle d'innovation Edtech sur le modèle des startups d'état aurait pour missions de :

- Piloter la mise en place d'outils permettant une large captation de données de comportements d'élèves face aux contenus pédagogiques numériques (vocabulaire, historique de travaux, programmes, etc.) pour permettre le développement d'algorithmes d'apprentissage réalisant cet accompagnement
- Définir et animer un laboratoire in vivo d'expérimentation d'outils pédagogiques reposant sur un réseau d'enseignants pilotes volontaires qui pourraient bénéficier d'allégements de classe, participer à des ateliers de co-développement avec les industriels et à des expérimentations avec des startups Edtech. Elle peut s'appuyer sur l'aide des délégués académiques au numérique (DAN) pour identifier des classes pilotes
- Définir et piloter un appel d'offre publique visant le développement de nouveaux outils par des acteurs ayant déjà développé des solutions autour de ces thèmes (Duolinguo, Lalilo, etc,)
    - il est à noter que les informations agrégées et non agrégées acquises dans ces déploiements devront être rendues disponible à l'Education Nationale pour le développement de futurs outils et l'analyse de l'impact des programmes

<u>Recommandation : Former les fonctionnaires et créer une expertise publique pour favoriser l'adoption de l'IA dans les administrations</u>

- L'ensemble des écoles de la fonction publique doit intégrer, à minima, un module obligatoire sur les risques et opportunités des I.A pour l'Etat et intégrant le sujet de la gouvernance par la donnée[18]
- Une expertise technique publique doit être développée sur l'intelligence artificielle au sein d'un des corps techniques de l'Etat (corps des mines et/ou le corps des administrateurs de

[17] Ce site permet la résolution de data challenge comme Kaggle mais est basée sur l'échange et la collaboration accélérant ainsi le processus et contribuant à la formation des participants. http://ramp.studio/
[18] « Gouverner à l'ère du Big Data » Elisabeth Grosdhomme Lulin

l'INSEE). Cette formation devra permettre à ses fonctionnaires de comprendre, définir et piloter des projets techniques et métiers s'appuyant sur une IA au sein des ministères

- Sur le modèle de Data.gouv.fr, une base structurée et labellisée de données doit être utilisée pour permettre la mise en œuvre par l'Etat la mise en œuvre de méthode d'apprentissage
- Une sensibilisation des élus est indispensable (parlementaires notamment)

Annexe 1 – Contributeurs

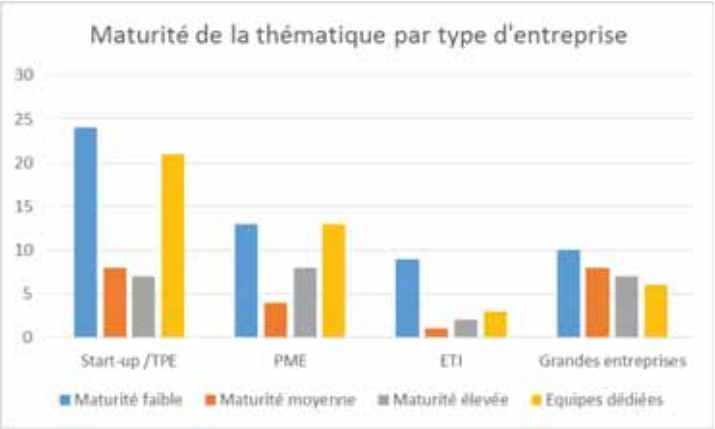
Catégorie	N	Nom	Prénom	Fonction & Affiliation
ACAD	FR	Vayatis	Nicolas	Prof. ENS-Paris-Saclay
TPE	FR	Grux	Alexandre	CEO, Odyssey38
	FR	Bazeille	Thomas	
ACAD	FR	Bar-Hen	Avner	Prof. CNAM
SSA	FR	Buffat	Stéphane	Chercheur IRBA
ACAD	FR	Chalmond	Bernard	Prof. Cergy
ACAD	FR	Graffigne	Christine	Prof. Paris-Descartes
ACAD	FR	Gramfort	Alexandre	MCF Telecom ParisTech
TPE	FR	Grondin	Marjolaine	CEO, JAM
ACAD	FR	Ralaivola	Liva	Prof. Aix-Marseille
TPE	US	Renard	Grégory	CTO, Xbrain
ETI	FR	Duranton	Marc	CEA List
PME	FR	Kaeser	Alain	CSO, Yseop
ETI	FR	Sapina	Nino	ex- CTO Software, Aldebaran
ACAD/TPE	FR	Saurel	Pierre	MCF Université Paris Sorbonne, Avocat
ACAD/ETI	FR	Milano	José	DGA KEDGE Business School
MINESR	FR	Goudon	Thierry	DGRI
DGE	FR	Anki-Zuccarello	Gisèle	Chargé de mission DGE

Annexe 2 – Analyse des interviews et des réponses au questionnaire

Les membres de notre groupe ont effectué une cinquantaine d'interviews longues avec des experts du secteur privé sur la question de l'implémentation de l'IA dans leur développement stratégique et nous avons réuni plus de 150 réponses à notre questionnaire.

Les membres du groupe de travail remercient Partech Ventures, ISAI et le pôle de Compétitivité Finance et Innovation pour l'aide apportée dans la diffusion du questionnaire.

1. **Maturité des entreprises**



Les startups sont les premiers acteurs de l'IA en France. Plus de 200 d'entre elles intègrent ou conçoivent des systèmes d'IA. Elles ont une capacité de recrutement limitée et des délais courts pour développer leurs produits (viabilité technologique est de 18 à 36 mois). Elles ont donc besoin de personnes à forte valeur ajoutée pouvant s'adapter rapidement et sont les premiers consommateurs d'expertise dans ce domaine.

Les PME/ETI semblent en majorité être encore en train de réaliser leur transition vers le numérique et ne se sont pas encore pleinement saisi de ces sujets.

Les grandes entreprises cherchent à accélérer l'automatisation de leur processus interne et expérimentent les possibilités offertes par les dernières technologies. Cela dit, une grande partie d'entre elles sont surtout sensibles à un effet de mode. Dans ces entreprises, les décideurs ne sont pas capable d'estimer l'impact réel qu'elles peuvent avoir ni la valeur qu'elles peuvent créer ou non, pour leur société et sur quel type de projet. Selon leur désir de maîtriser ces technologies, elles se tournent donc vers les solutions clés-en-main des géants de l'informatique (IBM, Microsoft, etc) ou vers les briques technologiques portées par des startups. Le manque d'information autour de ces technologies laisse la place à une influence forte des lsociétés influenceuses qui ont leur propre agenda.

Dans la plupart des entreprises sondées, y compris celles peu matures sur le sujet, quelques briques spécifiques d'IA au moins sont intégrées dans les processus métier et/ou leurs systèmes de marketing/relation client. Les entreprises disposant d'équipes spécialisées restent celles du secteur IT.

En chiffres :

- 40 entreprises répondantes n'utilisent pas l'IA, 50 ont des spécialistes l'utilisant au sein de leurs équipes métier, 15 n'ont pas d'équipe spécifiquement en charge et 45 ont des équipes dédiées.
- Les 113 entreprises interrogées utilisant l'IA utilisent en moyenne au moins 2 technologies (voire tableau ci-après)

2. Impact anticipé de l'IA sur les métiers et les besoins en formation

Tous les experts proches du domaine attendent un impact massif sur les organisations, entreprises ou administrations, et sur certaines fonctions au sein de celles-ci. Les principaux métiers (et compétence) qui devraient apparaître ou se généraliser sont :

- Architecte en conception d'IA : nécessitant une vue générale des techniques d'IA et une capacité à les composer, cela restera un poste d'expert mais les enjeux de compétitivité et de prévention du risque d'intermédiation devraient créer une forte demande pour ces profils
- Intégrateurs d'IA : développeurs informatiques avec une compréhension des techniques de machine et de deep learning qui adapteront ces briques technologiques à des usages particuliers
- Spécialistes « métier » : ayant une compréhension d'un métier et des données associées, qui soit capables d'organiser le problème à résoudre pour le rendre accessible aux solutions techniques existantes, de sélectionner et utiliser les nouveaux outils embarquant une IA et parfois d'entraîner et de superviser une IA.
- Designers d'interactions avec IA/robots : spécialiste de l'utilisation de données comportementale et de l'ergonomie pour concevoir et spécialiser les interfaces avec les utilisateurs moins qualifiés et les clients.
- Entraineurs d'IA : moins ou pas qualifiés sur les techniques d'IA mais ayant une haute spécialité de leur métier et qui formeront

On attend également une forte demande de chefs de projets intervenant de manière transversale sur le développement, l'intégration, et la maintenance des systèmes d'IA.

Machine learning	Système expert	Traitement du langage naturel	Robotique
75	47	52	26

Type d'IA utilisée par les entreprises interrogées

En chiffres :

- 75% des entreprises interrogées dans le cadre de notre étude estiment que l'intelligence artificielle aura un impact positif sur les emplois dans leur entreprise et seulement 7% négatif.[19]
- Sur les 45 entreprises ayant une équipe dédiée à l'IA, 1/3 recrutent leurs membres après un doctorat et les autre à BAC+5 (écoles d'ingénieurs et masters informatiques)
- Seules 20 entreprises seulement déclarent avoir mis en place des formations spécifiques pour leurs équipes concernées (exclusivement des PME/startups)
- 63 entreprises déclarent avoir prévu un plan de formation sur ce sujet (12 grands groupes, 6 ETI, 22 PME et 23 startups)

[19] 80% des entreprises déclarent que les employés impactés par l'adoption de l'IA seront redéployés à des postes plus qualifiés. (« AI Maturity Index» Infosys 2016)

Annexe 3 – Analyse des formations françaises en IA

Des formations de disciplines très différentes sont reliées à l'IA. On peut distinguer celles qui sont plutôt en amont de la création d'une "Intelligence Artificielle", comme par exemple celles qui touchent à l'organisation et la gestion des données ou à l'infrastructure matérielle (capteurs, technologies de calcul). De manière beaucoup plus générale, des étudiants issus de formations en sciences cognitives, en ergonomie, en éthique peuvent également contribuer de manière fondamentale à la création d'une IA. Sans amoindrir l'importance de ces formations qui ont de toute manière leur existence propre, il n'est évidemment pas envisageable de lister de façon exhaustive les possibilités. Citons juste, en illustration, le Cogmaster[20], le master Ergonomie et ingénierie facteurs humains[21], le master Spécialité Ethique de la santé et droits de la personne[22].

En aval, de nombreuses disciplines seront certainement amenées à utiliser, à divers degrés, des IA. Au sein même des disciplines scientifiques, on peut trouver par exemple des formations en robotique avec des applications variées telles que l'aide aux personnes dépendantes[23] et les services à la personne[24], mais aussi l'intelligence ambiante, les systèmes cyber-physiques[25] ou encore les applications nomades[26]. En sciences humaines et sociales (SHS), on trouve des formations orientées vers la fouille de données de types propres à la discipline[27], ou des travaux sur l'acceptabilité sociale de l'IA. En sciences cognitives on peut donner comme exemple le traitement automatique de la parole[28]. En santé de nombreuses applications sont en cours d'élaboration (cf. document GT IA et Santé) sans pour autant donner lieu à des formations adaptées lors des études de médecine. Pour certaines disciplines, telles que le Droit, les pistes d'utilisations potentielles commencent à se dessiner (GT IA et Droit) mais nous n'avons encore trouvé aucune formation dans ce créneau.

Les formations en IA

De manière contradictoire, il est plus compliqué de définir ce que l'on peut considérer comme une formation en IA que de discuter de ce que sont les applications de l'IA. Le choix fait ici consiste à décrire les formations qui vont former les "développeurs" d'Intelligences Artificielles au sens "cœur de métier" et donc avec une forte composante de mise en œuvre informatique ou de travaux de recherche sur les nouveaux objets ou concepts associés.

Regardons pour commencer les formations en traitement d'images. Étudié depuis relativement longtemps, prétraitement dans de nombreuses applications et ayant bénéficié de l'apport récent et très efficace du « deep learning », il est raisonnable de considérer le traitement d'images comme relevant de l'IA. Mais, simultanément, la grande majorité des formations en traitement d'images, que ce soit en écoles d'ingénieurs ou en universités, ne se focalisent pas sur les techniques aboutissant à une décision « artificielle » mais plutôt sur des traitements qui peuvent être utilisés dans des applications «

---

20 Cogmaster, Master recherche en sciences cognitive, http://sapience.dec.ens.fr/cogmaster/www/

21 Master Ergonomie et ingéniere facteurs humains, Université Paris Descartes, http://www.biomedicale.parisdescartes.fr/Ergonomie-et-Ingenierie-Facteurs.html?lang=fr

22 Spécialité Ethique de la santé et droits de la personne, Université de Lorraine, https://formations.univ-lorraine.fr/fr-FR/fiche/presentation/UL-PROG4185/UL-PROG4189

23 Master Ingénierie des Systèmes Complexes, Systèmes cyber-physiques, technologies de l'Information, de l'intelligence et du contrôle, Université Paris-Est Créteil, http://www.lab.lissi.fr/sctiic/start

24 Spécialité Robotique, assistance et mobilité, Université de Versailles Saint-Quentin en Yvelines, Université Paris-Saclay, https://www.universite-paris-saclay.fr/fr/education/master/m2-robotique-assistance-et-mobilite-ram#presentation-m2

25 Master Informatique, parcours Cyber-Physical Social Systems (CSP2), Université Jean Monnet, Saint-Etienne, https://depinfo.univ-st-etienne.fr/Master/cps2.php

26 Master Informatique, parcours informatique nomade, intelligence et sécurité, Université d'Orléans, http://formation.univ-orleans.fr/fr/formation/offre-de-formation/master-lmd-XB/sciences-technologies-sante-STS/master-informatique-specialite-informatique-nomade-intelligence-et-securite-finalite-professionnelle-et-recherche-program-scimif2-502-2.html

27 Master Informatique et Mathématiques Appliquées au Sciences Humaines et Sociales - MIASHS, Université Paul Valéry, Montpellier 3, http://ufr6.univ-montp3.fr/index.php/masters/133-master-miashs/170-programme-de-la-formation

28 Master Sciences de la cognition et application, parcours TAL, Université de Lorraine, https://formations.univ-lorraine.fr/fr-FR/fiche/presentation/UL-PROG1894/UL-PROG1897#debut

classiques » ou comme éléments de décision pour une IA.

Logiques non formelles, satisfaction de contraintes, web sémantique, logiques de description, systèmes multi-agents sont des intitulés classiques de grands domaines de l'IA. Il est bien rare de ne pas trouver au moins un enseignement de ce type dans n'importe quel parcours de master d'informatique. Avec des contenus relativement comparables, certaines descriptions de masters se positionne très fortement dans la thématique IA et d'autres beaucoup moins clairement. Nous dirons que ces masters ou parcours de master sont à dominante « IA généraliste ». Disposant d'un certain recul et d'une assise auprès des utilisateurs industriels, on trouve ici le panel complet des masters avec des parcours très professionnalisants, et même en apprentissage, par exemple le parcours Informatique décisionnelle[29], et d'autres qui débouchent aussi bien sur des thèses, des thèses CIFRE ou une embauche directe en entreprise. Plusieurs exemples déjà cités relèvent de cette dominante[4, 6, 7]. La prise en compte de l'aléatoire n'est pas ignorée dans ces formations, comme par exemple dans le master ANDROID[30], mais ce n'est pas l'élément qui sous-tend l'essentiel de la formation. On trouve aussi des parcours dédiés aux applications en robotiques, avec des composantes de traitement d'images et de la parole[31]. Les effets de mode et l'évidente qualité des résultats obtenus par certaines méthodes récentes très en vue actuellement ne doivent pas faire oublier la recherche très active dans ce domaine de l'IA et la nécessité d'un cadre général qui permet d'intégrer des informations de types très variés qui peut s'avérer utile dans le futur en lui associant les techniques plus récentes.

De manière similaire, les approches de type data mining, data science, machine learning, apprentissage, big data, deep-learning sont des mots clés classiques d'un autre type de formations en IA que nous appellerons à dominante "apprentissage". Ces formations requièrent des étudiants un bagage en informatique, en statistique et en mathématiques appliquées. Elles sont parfois rattachées à l'informatique, parfois aux mathématiques ou encore aux deux disciplines. Les étudiants formés acquièrent des savoirs qui sont par exemple très prisés par les entreprises qui doivent extraire de la connaissance de grands volumes de données (Google, Facebook, Amazon…). On trouve dans cette dominante des masters très sélectifs tels que ceux de l'ENS Cachan[32] ou de l'université Pierre et Marie Curie[33]. D'autres masters permettent d'illustrer différentes spécificités telles qu'une volonté forte d'ouverture vers l'international[34], ou une spécialisation vers les techniques d'apprentissage automatique et le traitement automatique de la langue[35] ou encore sur l'articulation entre apprentissage automatique et recherche de l'information[36]. La dynamique relativement récente du Big Data, et plus précisément de la science des données, mais aussi la réalité du terrain qui confronte des intervenants très variés à la multiplication des données a conduit à l'apparition récente de nombreuses formations de ce type[37]. C'est dans ce domaine que l'on trouve les formations les plus professionnalisantes et le seul domaine dans lequel nous avons identifié des licences professionnelles, telles que celle de Grenoble[38].

Enfin, certains masters mêlent les deux dominantes définies ci-dessus (i.e. « IA généraliste » d'une

---

[29] Master Miage en alternance, parcours Informatique décisionnelle, Université Paris Dauphine, http://formations.dauphine.fr/offre/fr-FR/fiche/A5STI/programme//

[30] Master ANDROIDE: AgeNts Distribués, Robotique, Recherche Opérationnelle, Interaction, DEcision Université Pierre et Marie Curie, http://www-master.ufr-info-p6.jussieu.fr/lmd/specialite/androide/

[31] Master Intelligence Artificielle et Reconnaissance des Formes, Université de Toulouse, Toulouse 3, https://masteriarf.wordpress.com/

[32] Master MVA- Mathématiques, vision, apprentissage, ENS Cachan, www.math.ens-cachan.fr/version-francaise/formations/master-mva/

[33] Master DAC - Données apprentissage et connaissances, Université Pierre et Marie Curie, http://dac.lip6.fr/master/

[34] Master international Machine Learning and Data Mining, Université Jean-Monnet, Saint-Etienne, http://mldm.univ-st-etienne.fr/overview.php

[35] Master Apprentissage et Traitement Automatique de la Langue (ATAL), Université de Nantes et Université du Mans, Université du Maine, http://www.master-info.univ-nantes.fr/00542841/0/fiche___pagelibre/&RH=1403710895111 ou http://www-info.univ-lemans.fr/?page_id=10

[36] M2 Apprentissage, Information et Contenu - AIC, Université Paris-Sud, Paris-Saclay, https://www.universite-paris-saclay.fr/fr/education/master/m2-apprentissage-information-et-contenu-machine-learning-information-and-content#presentation-m2

[37] M2 Apprentissage automatique et données massives (machine learning & data science), Université Lille 3, https://www.univ-lille3.fr/ufr-mime/formation/master-mdls/

[38] Licence pro Big data, Université Grenoble-Alpes, IUT2, http://stid-grenoble.xtek.fr/1399_0_0_0_LP+Big+Data_0.html

part et IA de type « apprentissage », d'autre part) ou portent plusieurs parcours avec des cours communs[39] et pourraient peut-être contribuer à préparer leurs étudiants à développer des techniques mixtes.

Les Grandes Ecoles ont le plus souvent construit durablement des collaborations et des masters communs avec leur environnement universitaire. De nombreuses écoles d'ingénieurs développent des partenariats avec les Universités voisines. Dans tous les cas, cela a un effet bénéfique de part et d'autre, en particulier en termes de bassin de recrutement comme indiqué plus loin. Les écoles d'ingénieurs ont plus de facilité que les Universités pour mettre en place rapidement des cursus en phase avec l'actualité et c'est peut-être une capacité qui pourrait être mieux valorisée vis à vis des interactions avec le monde universitaire.

Si les formations de type master sont les plus visibles, sur un domaine aussi pointu que l'IA, il faut évidemment aussi prendre en compte la formation doctorale. Les résultats d'une simple recherche sur le web des mots clés "formation doctorale" ou "doctorat" et "intelligence artificielle" ou même simplement "informatique" prête à réflexion (mais ça n'est pas spécifique à la discipline). Le même genre de recherche avec des mots clés identiques en anglais doit alerter sur l'attractivité au niveau international et même national, en effet on voit immédiatement apparaître dans ce cas des listes de formations sélectionnées et de grandes universités anglo-saxonnes. Les masters considérés ci-dessus sont systématiquement associés à un ou des laboratoires de recherche de qualité, tous débouchent sur la possibilité de thèses, avec un encadrement lui aussi de qualité. La structuration en écoles doctorales, souvent pluridisciplinaires, a un impact fort sur la visibilité de ces formations. La nomenclature des mentions de master et la disparition des spécialités aura probablement un impact là aussi en termes de visibilité.

Types de formation

Un autre point d'entrée pour distinguer les formations est celui des types de formations. On distingue les parcours/spécialités pro ou recherche – il faut noter que la plupart des formations répertoriées n'affichent pas nécessairement une distinction claire entre parcours pro et recherche – et les parcours en alternance, en faible nombre.

Nous pouvons noter l'absence de formations continues. Le risque est que ce créneau soit préempté par des entreprises spécialisées sur l'usage (et non la création) de technologies d'intelligence artificielle très opportunistes (cf. le cas du Deep Learning actuellement : il existe des utilisateurs très compétents de Deep Learning dans les entreprises, utilisateurs qui pourraient dispenser des formations essentiellement technologiques sur le sujet).

Il serait judicieux qu'en plus de l'essor des formations initiales en intelligence artificielle se développent plus de formations en alternance sur le sujet et que des mécanismes pilotés par les universités rendent les enseignants-chercheurs du domaine les interlocuteurs prioritaires pour les formations continues et celles dispensées en entreprise.

Bassin de recrutement

Enfin, et c'est un constat classique pour toutes les formations en France, on note une réelle disparité d'attractivité entre les formations parisiennes et les formations qui se trouvent en région. En particulier, deux formations, le master MVA et le master DAC attirent des étudiants ayant suivi des

[39] Master Informatique, spécialités Intelligence artificielle et Data Science, Université Lyon 1, http://master-info.univ-lyon1.fr/IA/ et http://master-info.univ-lyon1.fr/DS/

parcours académiques de tout premier plan et sont obligés chaque année de refuser de brillants étudiants.

Les Grandes écoles et les écoles d'ingénieurs disposent d'un vivier d'étudiants qui ont les compétences techniques, qui sont naturellement directement concernés par l'évolution technologique et qui constitue donc un levier évident de transfert technologique. En termes de répartition géographique, cela induit néanmoins des disparités qui peuvent être notables.

Au niveau de la formation doctorale, le constat généralisé actuellement est que le nombre de bourses de thèses ne suffit pas financer les étudiants qui le mériteraient ni à développer l'ensemble des sujets de thèses pertinents. Des bourses CIFRE existent, avec des disparités géographiques bien évidemment, mais on constate que, et ça n'est pas spécifique à l'IA, le nombre d'entreprises qui ont recours à ces bourses est relativement faible et certainement bien moindre que ce qu'on pourrait envisager dans un domaine aussi porteur. Néanmoins, l'ANRT joue pleinement son rôle à ce niveau et les entreprises n'ont peut être pas encore majoritairement identifié leurs besoins dans la thématique IA.

Le développement de masters internationaux sur la thématique de l'intelligence artificielle mériterait d'être accentué.

Annexe 4 – Recommandation 2.1 développée

Recommandation 2.1 : Instauration d'un enseignement "IA, traitement des données et sciences numériques" de l'école primaire au lycée

- Un enseignement « IA, traitement des données et sciences numériques » au primaire.
- Identifier les compétences « IA, traitement des données et sciences numériques » des enseignants du secondaire et exploiter cet élément pour mener une politique d'affectation sur des postes de référent numérique dans chaque établissement.
- Impliquer les enseignants de toutes disciplines pour un usage critique des outils de l'IA.
- Renforcer les options ouvrant vers les métiers du numérique
- Ouvrir et développer une plate forme mutualisée d'outils pédagogiques.

Cette recommandation vise à instaurer ou renforcer un enseignement « IA, traitement des données et sciences numériques » de l'école au lycée. Cet enseignement, critique, multi-forme et pluri-disciplinaire, doit montrer que les aspects techniques sont inter-dépendants des enjeux éthiques, sociétaux et des applications. Cet effort sera naturellement amené à diffuser dans toute la population.Il s'adosse aux apprentissages en mathématiques et implique l'utilisation d'applications et une initiation à la programmation.

La priorité consiste à instaurer un tel enseignement pour le **primaire en l'identifiant symboliquement en tant que discipline**. Cet enseignement s'adosserait naturellement aux apprentissages en mathématiques (classement, regroupement et traitement de données, travail sur des procédures systématiques, traitement de problèmes par des outils d'intelligence artificielle : affectation de ressources, optimisation de flux...). Comme cela est désormais prévu dans les programmes, on peut aborder l'initiation à la programmation, notamment via un outil comme scratch ou via l'usage de robots comme les bee-bots. Cela impose une évolution des programmes du premier degré surtout pour que les enseignements au numérique déjà présents soient facilement identifiables en étant regroupés sous un unique intitulé disciplinaire, et requiert un minimum de formation du personnel enseignant, ainsi qu'un investissement matériel (tablettes, robots...).

Au collège et au lycée, il existe déjà un certain nombre d'enseignements qui permettent d'aborder les problématiques des sciences numériques. L'utilisation de scratch est vue comme un succès qui permet de structurer une certaine rigueur et développe des qualités intéressantes (persévérance, autonomie...). Un enjeu consiste certainement à impliquer les enseignants de disciplines non scientifiques dans l'utilisation d'outils numériques et le développement d'un regard critique sur cette utilisation (questions éthiques, juridiques, de sécurité,...). Une recommandation serait aussi de renforcer et faire monter en puissance les options ISN-ICN pour permettre une ouverture aux formations supérieures et aux métiers du numérique et de l'IA.

A l'heure actuelle, ces enseignements du second degré sont principalement liés aux cours de mathématiques et de technologie. Une faiblesse réside en partie dans la formation inégale des enseignants. La formation continue qualifiante est bien adaptée et des options de divers concours de recrutement d'enseignants sont très clairement identifiées sur les sciences informatiques (par exemple, aux concours de l'agrégation il y a une option "informatique et système d'information" aux concours économie-gestion, il y a une option "ingénierie informatique" en sciences industrielles de l'ingénieur et une option informatique en mathématiques). Il est donc nécessaire d'identifier ces enseignants ayant ces compétences spécifiques, par exemple par un code de sous discipline, et d'utiliser cet élément comme un argument d'affectation afin de garantir une répartition des compétences équilibrée et répondant aux besoins des établissements. Nous proposons en effet de créer des postes spécifiques de « **référent numérique** » au sein des établissements (collèges, lycées) en charge de la coordination et de l'animation des enseignements « IA, traitement des données et sciences numériques ». A terme, il faudra probablement aussi s'interroger sur l'élargissement du vivier des candidats aux concours

d'enseignement et chercher à attirer des profils plus spécialisés sur les sciences informatiques.

Une recommandation complémentaire consiste à favoriser l'émergence et la diffusion d'un **corpus de contenus d'enseignement** couvrant tous les niveaux en algorithmique, informatique et intelligence artificielle. A cette fin, un portail national pourrait être ouvert, mettant à disposition ces contenus de manière libre et ouverte (pour toute la communauté éducative : enseignants, élèves, parents,...). Le développement d'une telle plate-forme nécessite un travail de coordination, mais doit reposer sur la valorisation d'initiatives, par exemple sous forme de décharges ponctuelles d'enseignement accordées sur évaluation de projets.

Annexe 5.1 – Fiche de formation, module Initiation à l'IA

Intelligence Artificielle : Initiation

Durée : 7h (ou 14h ?)

Public : tous

Objectifs : fournir un éclairage et sensibilisation à l'Intelligence Artificielle pour tous : Historique / divers types d'IA / les enjeux / les mythes et réalités / le futur de l'IA

L'objectif de ce module est de donner aux participants une meilleure compréhension de ce qu'est réellement l'IA, au-delà des annonces médiatiques (que le module abordera tout de même) et des effets de mode et donc inévitables confusions et amalgames, en fournissant une information plus approfondie qu'une information de type journalistique..

Programme :

1. **Historique**
   - Pourquoi l'IA ? Un vieux rêve de l'humanité.
   - Du General Problem Solver à AlphaGo, des système experts au machine learning
     - De la conférence de Dartmouth College de 1956 aux principes d'Asilomar de 2017
     - Phases de dépression de l'IA puis de renouveau
     - l'IA de nos jours. Quelques évènements marquants. Positionnement des grands acteurs économiques

2. **Les divers types d'IA**
   - IA faible, IA forte
   - IA symbolique/logique vs IA probabiliste/numérique
   - Liens entre l'analyse de données (*data analytics*, Big Data) et l'IA
   - Traitement du langage naturel
   - Compréhension de scènes
   - Robotique

3. **Cas d'application**
   - Quelques coups d'éclat grand public
   - Dans la vie de tous les jours : de l'assistant personnel au véhicule autonome
   - Dans le monde de l'entreprise : Industrie, services

4. **Les enjeux de l'IA**
   - Les impacts sociétaux, économiques, politiques et éthiques
     - Impacts futurs : peurs et fantasmes, la singularité technologique
     - Le contexte international et les différentes attitudes face à l'IA
     - Impacts maintenant et dans le futur proche
   - Un peu de prospective …
   - Comment se positionner maintenant, en tant que citoyen et/ou en tant qu'acteur de la vie économique ?

Annexe 5.2 – Fiche de formation, module IA théorie et pratique

Intelligence Artificielle : théorie et pratique

Durée : 7 h

Prérequis : avoir suivi le module Initiation

Public : Pour les curieux voulant plus qu'une introduction de base sur le fonctionnement des IA, pour les futurs responsables de projets IA. Accessible à des non informaticiens.

Objectifs : accéder à des exemples concrets de mise en œuvre d'IA de divers types. Effectuer des travaux pratiques de paramétrage / supervision d'IAs.

Ce module met en œuvre :

- des travaux pratiques sur ordinateur, sur la base d'applications prédéfinies et que l'élève sera amené à modifier (à la marge). On leur réservera au minimum 4h (donc 2h par cas). On peut envisager que les 2 cas pratiques soient similaires (par exemple mise en place d'un agent conversationnel ou reconnaissance de formes) mais avec deux approches différentes
- précédés de présentations formelles qui complètent donc la partie « Les divers types d'IA » du module d'initiation

Programme :

1. **IA probabiliste : principes de fonctionnement**
   - Apprentissage supervisé
   - Apprentissage non supervisé – Deep learning
   - Apprentissage par renforcement

2. **IA probabiliste : cas pratique**
   - Supervision d'une application faisant appel à un algorithme d'apprentissage supervisé

3. **IA symbolique : principes de fonctionnement**
   - Logique propositionnelle
   - Logique du premier ordre
   - Extensions (en très bref) : Logique modale, Logique temporelle

4. **IA symbolique : cas pratique**
   - Paramétrage (ajout de règles) d'un système d'aide à la décision ou agent conversationnel

5. **Comparaison des approches probabilistes et symboliques**

Annexe 5.3 – Fiche de formation, module "IA dans l'entreprise"

Intelligence Artificielle : l'IA dans l'entreprise

Durée : 7 h

Prérequis : avoir suivi le module Initiation

Public : Managers, acheteurs, responsables de projet, créateurs d'entreprises, DRH

Objectifs : exemples de réussites et d'échec de l'IA, les prérequis, exemples de coûts, impacts sur l'entrepris.

Ce module donne des clés de compréhension et de décision pour les acteurs en entreprise se trouvant en position :

- D'acteur dans une entreprise créateur d'une solution d'IA,
- D'acheteur d'une solution toute faite d'IA ou d'une plate-forme technologique IA

Programme :

1. **Pourquoi l'IA au sein de l'entreprise ?**
   - Augmenter la valeur :
     - Créer de nouveaux services
     - Etendre des services existants
   - Optimiser les processus et réduire les coûts

2. **Champs d'application de l'IA**
   - Illustration de cas d'usages (assortis d'exemples d'applications opérationnelles) par grand type d'activité
     - Services
     - Analyse de la production
     - Retail, relations avec le client
     - Autres
   - Et par métiers :
     - Vente
     - Production
     - Reporting
   - Usages futurs court et moyen terme

3. **Lancer un projet d'IA**
   - Clés de succès
     - Ne pas se tromper sur les objectifs
     - L'équipe interne
     - Les fournisseurs de technologie IA
       - Qualité de la technologie
       - Adaptation du type de technologie au cas traité
       - Qualité de l'accompagnement et des savoir-faire
   - Organisation projet
   - Délais et coûts

- Illustration sur quelques cas d’usage

4. Impacts de l’IA sur l’entreprise

- Nouveaux métiers, évolution de métiers existants
- Accompagnement au changement des collaborateurs … et des clients
- Impact sur l'organisation de l'entreprise

Annexe 6.1 – Groupe de référence : Contacts Interviewés

Entreprise	Fonctions	Type[40]	Prénom	Nom
ATOS-Wordline		Grande entreprise	Frédéric	Oblé
AXA France	Directrice du recrutement	Grande entreprise	Frédérique	Bouvier
CEA LETI	Directeur du Service Technologies de la Détection	Organisme de recherche	Jean-Marc	Dinten
DENTONS	Directrice de l'innovation	Grande entreprise	Marie Elisabeth	Bernard
ENA	Directrice de la formation	Ecole de la fonction publique	Nathalie	Tournyol
Holberton School	Fondateur	Start-up	Julien	Barbier
IBM France	Vice-Président	Grande entreprise	Silvano	Sansoni
IBM Watson	Directeur de la practice cognitive	Grande entreprise	Emmanuel	Vignon
Institut des Actuaires	Présidente de la Commission scientifique	Association professionnelle	Florence	Picard
Keras	Fondateur	Opensource	François	Chollet
Lagardère	CDO	Grand entreprise	Olivier	Sorba
Malakoff Médéric	Head of Digital transformation	Grande entreprise	Ivan	Rignault
MENESR	IPR Académie de Nantes	Fonction publique	Sandrine	Fleurant

[40] Typologie par nombre d'employés : Start-up, moins de 10 employés ; PME, entre 10 et 250 employés; Grande entreprise, plus de 250 employés

MGEN	Directeur Relation Client	Grande entreprise	Samuel	Pichot
Ministère de la Justice	Président TGI Strasbourg	Fonction publique	Thierry	Ghera
OpenClassroom	Fondateur	PME	Mathieu	Nebra
Panasonic	Data scientist	Grande entreprise	Grégory	Senay
Quantmetry	Fondateur	PME	Jeremy	Harroch
Recherche	Postdoc	Académique	Mathieu	Cisel
Regaind	CEO	Start-up	Arnaud	Laurenty
Renault		Grande entreprise		
Safran Aircraft Engines (SNECMA)	Expert in algorithms	Grande entreprise	Jérome	Lacaille
SFR	Directeur Relation Client	Grande entreprise	Pascal	Landré
SoftBankRobotics	CSO	ETI	Rodolphe	Gelin
Spirops	CEO	PME	Axel	Buendia
Spoon	CEO	Start-up	Jerome	Monceaux
Télécom ParisTech	Dir. de la Recherche	Académique	Patrick	Duvaut
Thalès		Grande entreprise		
Tharios	Fondateur	Startup	Christophe	Bourgoin

Annexe 6.2 – Groupe de reference: Entreprises sondées

Entreprise	Secteur d'activité	Type[41]
2B UP	Banque - Finance - Assurance	Start-up
Ailancy	Banque - Finance - Assurance	PME
Amdocs	Electronique - Informatique	Grande entreprise
ANEO	Electronique - Informatique	PME
Assureur Français	Banque - Finance - Assurance	ETI
AsVoR	Communication - Information	PME
Aurel BGC	Banque - Finance - Assurance	ETI
B-PART SAS	Banque - Finance - Assurance	Start-up
BA consulting	Banque - Finance - Assurance	Start-up
Banque Française Mutualiste	Banque - Finance - Assurance	ETI
BEEAM	Banque - Finance - Assurance	Start-up
Bizkitt AI	Electronique - Informatique	Start-up
Blueway	Electronique - Informatique	PME
BNP Paribas	Banque - Finance - Assurance	Grande entreprise
BNP PARIBAS	Banque - Finance - Assurance	Grande entreprise
BNP Paribas	Banque - Finance - Assurance	Grande entreprise
Bnp paribas	Banque - Finance - Assurance	Grande entreprise
BNP Paribas	Banque - Finance - Assurance	Grande entreprise
BNP Paribas Securities Services	Banque - Finance - Assurance	Grande entreprise
BPSS	Banque - Finance - Assurance	Grande entreprise
Brennus Analytics	Electronique - Informatique	Start-up
Bress Healthcare	Santé	Start-up
CA STORE	Banque - Finance - Assurance	Start-up
CACEIS	Banque - Finance - Assurance	ETI
caisse des dépôts	Banque - Finance - Assurance	Grande entreprise
Cartes Bancaires CB	Banque - Finance - Assurance	PME
CashLab	Banque - Finance - Assurance	Start-up
CENTICH MUTUALITE FRANÇAISE	Santé	ETI

[41] Typologie par nombre d'employés : Start-up, moins de 10 employés ; PME, entre 10 et 250 employés ; Grande entreprise, plus de 250 employés

CFVC	Banque - Finance - Assurance	Start-up
CGI	Electronique - Informatique	Grande entreprise
Chappuis Halder	Banque - Finance - Assurance	PME
conseil d'Etat	Droit	ETI
Contrari	Droit	Start-up
Coopérative de propriétaires	Banque - Finance - Assurance	Start-up
Coorpacademy	Enseignement - Formation	PME
CREDIT agricole	Banque - Finance - Assurance	Grande entreprise
Crédit Agricole	Banque - Finance - Assurance	Grande entreprise
Crédit Agricole CIB	Banque - Finance - Assurance	Grande entreprise
Criteo	Numérique - Multimédia - Audiovisuel	ETI
CrossQuantum	Banque - Finance - Assurance	PME
CWB	Electronique - Informatique	Start-up
CyberSAS	Communication - Information	Start-up
Deepomatic	Electronique - Informatique	PME
Deepor	Santé	Start-up
Deloitte	Banque - Finance - Assurance	Grande entreprise
Dentons	Droit	Grande entreprise
Digital Airways	Electronique - Informatique	PME
Digital Management Consulting	Communication - Information	Start-up
DV Consulting	Electronique - Informatique	Start-up
Easytransac	Banque - Finance - Assurance	Start-up
EIFR	Enseignement - Formation	Start-up
Elaia Partners	Banque - Finance - Assurance	Start-up
Ellis-Car The Data Driven Company	Banque - Finance - Assurance	PME
engie résseaux	Environnement - Nature - Nettoyage	ETI
ESTIMEO	Banque - Finance - Assurance	Start-up
Exact	Electronique - Informatique	PME
Femtopay	Banque - Finance - Assurance	Start-up
Formee	Communication - Information	Start-up
Forsides	Banque - Finance - Assurance	PME
Fortia Financial Services	Banque - Finance - Assurance	PME
GEMA	Electronique - Informatique	Start-up

Georges.tech sas	Numérique - Multimédia - Audiovisuel	Start-up
Gien Automatismes	Electronique - Informatique	PME
Gleetr	Electronique - Informatique	Start-up
Goldman Sachs	Banque - Finance - Assurance	Grande entreprise
Google	Communication - Information	Grande entreprise
Groupe La Poste	Social - Services à la personne	Grande entreprise
Groupe SEB	Electronique - Informatique	Grande entreprise
healsy	Santé	Start-up
Hi Bruno	Banque - Finance - Assurance	Start-up
Horizon Software	Electronique - Informatique	PME
IAE Dijon	Enseignement - Formation	PME
Implicity	Santé	Start-up
ingies	Numérique - Multimédia - Audiovisuel	Start-up
Insidor	Gestion - Audit - Ressources humaines	Start-up
Institut Louis Bachelier	Banque - Finance - Assurance	PME
INTIs	Numérique - Multimédia - Audiovisuel	PME
IUT Mantes en Yvelines	Sciences - Maths - Physique	PME
Ivalua	Electronique - Informatique	PME
Jam	Numérique - Multimédia - Audiovisuel	Start-up
jj	Banque - Finance - Assurance	PME
Kernix	Electronique - Informatique	PME
La Banque Postale	Banque - Finance - Assurance	ETI
Lamsade	Enseignement - Formation	ETI
LBPAM	Banque - Finance - Assurance	PME
LCL	Banque - Finance - Assurance	Grande entreprise
LIBERTY SPECIALTY MARKETS	Banque - Finance - Assurance	ETI
LIFE PLUS	Santé	Start-up
Lingua Custodia	Numérique - Multimédia - Audiovisuel	Start-up
LISM	Industrie - Matériaux	PME
LMA	Gestion - Audit - Ressources humaines	Start-up
M Capital Partners	Banque - Finance - Assurance	PME
Macif	Banque - Finance - Assurance	Grande entreprise
MAIF	Banque - Finance - Assurance	Grande entreprise
Malakoff Médéric	Banque - Finance - Assurance	Grande entreprise

MGEN	Banque - Finance - Assurance	Grande entreprise
MINALEA	Banque - Finance - Assurance	Start-up
MondoBrain	Banque - Finance - Assurance	PME
Moodeen	Hôtellerie - Restauration - Tourisme	Start-up
moodsights	Banque - Finance - Assurance	Start-up
MoodSights	Banque - Finance - Assurance	Start-up
MoodSights	Numérique - Multimédia - Audiovisuel	Start-up
My Pharma Company	Banque - Finance - Assurance	Start-up
MyDataBall	Numérique - Multimédia - Audiovisuel	Start-up
Nalo	Banque - Finance - Assurance	Start-up
Napoleon Capital	Banque - Finance - Assurance	Start-up
Natural Scurity Alliance	Banque - Finance - Assurance	Start-up
Natural Security Alliance	Banque - Finance - Assurance	Start-up
Neiter consulting	Enseignement - Formation	Start-up
Nexialog	Banque - Finance - Assurance	PME
NEXYAD	Banque - Finance - Assurance	Start-up
Novaquark	Culture - Spectacle	PME
odase ontologies	Electronique - Informatique	Start-up
OFI Asset Management	Banque - Finance - Assurance	ETI
opus finance	Banque - Finance - Assurance	PME
Orange	Numérique - Multimédia - Audiovisuel	Grande entreprise
ORANGE SA	Banque - Finance - Assurance	Grande entreprise
Oyst	Commerce - Immobilier	PME
Panasonic	Numérique - Multimédia - Audiovisuel	Grande entreprise
Paris Region Entreprises	Electronique - Informatique	PME
PIM	Banque - Finance - Assurance	PME
Premaccess	Electronique - Informatique	Start-up
Premiumpeers	Banque - Finance - Assurance	Start-up
PW-Consultants	Banque - Finance - Assurance	PME
Quantmetry	Electronique - Informatique	PME
Regaind	Electronique - Informatique	Start-up
Regaind	Electronique - Informatique	Start-up
REMY &PARTENAIRES	Humanitaire	Start-up
Robin'Finance	Banque - Finance - Assurance	Start-up

Robin'Finance	Banque - Finance - Assurance	PME
SESAMm	Banque - Finance - Assurance	Start-up
SFR	Numérique - Multimédia - Audiovisuel	Grande entreprise
SG	Banque - Finance - Assurance	Grande entreprise
SIWAY	Communication - Information	PME
Société Générale Insurance	Banque - Finance - Assurance	ETI
SoftBank Roboticz	Numérique - Multimédia - Audiovisuel	ETI
Spoon	Electronique - Informatique	Start-up
SR2C Consulting & Managment	Banque - Finance - Assurance	Start-up
Steeple	Numérique - Multimédia - Audiovisuel	Start-up
Streamdata.io	Electronique - Informatique	PME
SysperTec	Electronique - Informatique	PME
TacoTax	Banque - Finance - Assurance	Start-up
Tomco	Commerce - Immobilier	Start-up
Ubisoft	Numérique - Multimédia - Audiovisuel	Grande entreprise
Université Paris Est Marne La Vallée	Enseignement - Formation	ETI
Walnut Algorithms	Banque - Finance - Assurance	PME
WELBHELP	Social - Services à la personne	Grande entreprise
WFC	Banque - Finance - Assurance	Start-up
xBrain	Numérique - Multimédia - Audiovisuel	PME
Xerox	Communication - Information	Grande entreprise

2. INDUSTRIALISER ET TRANSFÉRER LES TECHNOLOGIES DE L'IA VERS LES AUTRES SECTEURS ÉCONOMIQUES EN MAXIMISANT LES RETOMBÉES ÉCONOMIQUES SUR LE TERRITOIRE NATIONAL

GROUPE DE TRAVAIL 2.1
TRANSFERT DE TECHNOLOGIES

Groupe Industrialisation et appropriation des résultats de la recherche par les entreprises

1. Rapport du groupe de travail

1.1 Périmètre d'étude du groupe de travail

La mission confiée au groupe visait à :

1. Identifier les principaux domaines et sous domaines de la recherche en Intelligence artificielle dans lesquels un transfert à court terme (moins de trois ans) est envisageable vers les entreprises, en tenant compte autant que possible de l'effort d'industrialisation et du potentiel économique ;
2. Proposer des actions (notamment appels à projets) permettant de susciter et de soutenir les coopérations entre entreprises et laboratoires favorisant les transferts de technologies ;
3. Caractériser, au regard des sujets de recherche identifiés dans le sous-groupe 1, les phases de tests et d'essais qui seraient à réaliser, en identifiant autant que possible les risques réglementaires éventuels ;
4. Identifier ou proposer des dispositifs et moyens (plateformes matérielles et logicielles) possiblement mutualisés permettant aux entreprises de tester les technologies issues de travaux de recherche ;
5. Et toute autre action faisant sens sur le sujet de l'industrialisation des résultats de la recherche en IA.

Le groupe est composé de représentants du monde académique et d'entreprises (start-ups, PME et grands groupes), ainsi que de trois représentants de l'Etat (Direction Générale des Entreprises, Commissariat Général à l'Investissement et Secrétariat Général pour la Modernisation de l'Action Publique) et d'un membre du Laboratoire National d'Essais. Nous avons constitué un groupe principal ainsi que des sous-groupes thématiques dont l'animateur est membre du groupe principal.

1.2 Résumé exécutif

La France dispose d'une recherche de tout premier plan en intelligence artificielle (IA), notamment grâce à la qualité de son enseignement dans les disciplines de base indispensables à la maîtrise des technologies sous-jacentes : informatique, mathématiques pures et mathématiques appliquées, logique, formalisation des raisonnements et des connaissances. Ces disciplines sont essentielles pour l'ensemble des domaines de l'IA tels que l'apprentissage automatique, la perception, l'optimisation combinatoire, le raisonnement et l'aide à la décision, le traitement de la langue naturelle, l'interaction humain-machine, et le web sémantique.

Notre pays dispose également d'un tissu industriel susceptible de bénéficier des technologies d'intelligence artificielle dans de nombreux secteurs : énergie, santé, transport, défense, finance, numérique, pour n'en citer que quelques-uns.

La qualité de notre enseignement et de notre recherche en IA est reconnue internationalement, à tel point que de nombreuses entreprises internationales investissent dans notre pays pour y être

confrontées et pour y recruter des jeunes diplômés. Il est essentiel que les entreprises françaises puissent également bénéficier de cet atout, ce qui n'est pas encore suffisamment le cas.

Telle était donc la mission confiée au groupe « Industrialisation et appropriation des résultats de la recherche par les entreprises », détaillée comme suit :

1. Identifier les principaux domaines et sous domaines de la recherche en Intelligence artificielle dans lesquels un transfert à court terme (moins de trois ans) est envisageable vers les entreprises, en tenant compte autant que possible de l'effort d'industrialisation et du potentiel économique *;*
2. Proposer des actions (notamment appels à projets) permettant de susciter et de soutenir les coopérations entre entreprises et laboratoires favorisant *les transferts de technologies ;*
3. Caractériser, au regard des sujets de recherche identifiés, les phases de tests et d'essais qui seraient à réaliser, en identifiant autant que possible les risques réglementaires éventuels ;
4. Identifier ou proposer des dispositifs et moyens (plateformes matérielles et logicielles) possiblement mutualisés permettant aux entreprises de tester les technologies issues de travaux de recherche ;
5. Et toute autre action faisant sens sur le sujet de l'industrialisation des résultats de la recherche en IA.

Le groupe, composé à parts égales de membres de la communauté académique et de représentants du monde des industries et services, a questionné les équipes de recherche et les entreprises sur le potentiel de transfert, les enjeux et les freins à lever pour augmenter la bande passante entre ces deux communautés dans le domaine de l'intelligence artificielle et dans ses sous-domaines. Plus d'une centaine de technologies développées dans les laboratoires de recherche ont été jugées transférables à court et moyen terme vers les entreprises, essentiellement dans les sous-domaines de l'apprentissage automatique, de l'optimisation et de la prise de décision, du traitement de la langue naturelle, de la représentation des connaissances, de l'évolution artificielle et des systèmes multi-agents. En résonance avec ces résultats, pour la cinquantaine d'entreprises qui ont répondu au questionnaire, les bénéfices attendus des technologies d'IA sont qualifiés comme importants voire critiques.

Pour ce faire, les principaux enjeux identifiés par les chercheurs et les entreprises, qui correspondent également à des freins à lever, sont les suivants :

- **Disposer d'ensembles de données nettoyées et fiables prêtes à l'emploi**, notamment données publiques ou ouvertes. Elles sont nécessaires aux algorithmes d'apprentissage et plus largement pour tous les domaines qui font appel à des données, qu'elles soient expérimentales, acquises en ligne, traces de transactions, etc. L'application des méthodes d'IA sur des données représentatives des problèmes à traiter est une condition sine qua non de leur adoption par le monde économique.
- **Etre capable de garantir ou de certifier certaines méthodes** (besoin de fondements théoriques pour prouver les algorithmes). Cette garantie/certification est indispensable pour les systèmes critiques et importante pour les autres, faute de quoi les entreprises seront réticentes à déployer des systèmes auxquels elles ne font pas totalement confiance.

- **Pouvoir accéder à des compétences humaines opérationnelles pour les entreprises dans l'écosystème de recherche et d'enseignement supérieur** :
    - en recrutant des étudiants formés à l'IA (notamment via des bourses CIFRE) ;
    - en assurant la formation continue en IA des collaborateurs ;
    - en s'appuyant sur des intégrateurs de technologies ;
    - en offrant des conditions attractives pour enrayer la fuite des cerveaux.
- **Disposer d'infrastructures IA (données, plateformes) et de moyens de calcul au niveau des défis**, accessibles en France, qui sont également des enjeux de souveraineté nationale. Les moyens actuels ne sont pas adaptés aux besoins des systèmes d'IA. Il faut disposer de plateformes correspondant aux exigences des applications IA et capables de se positionner dans la compétition internationale.
- **Le besoin de maturation des technologies** issues de la recherche à un niveau de qualité et de performance exigé par les entreprises. Il n'est pas attendu des chercheurs qu'ils produisent des résultats directement exploitables par les entreprises, mais plutôt des démonstrateurs, preuves de concept, prototypes. Il est donc important de disposer de ressources capables de faire « maturer » les technologies pour assurer leur appropriation.
- Pouvoir répondre aux **exigences de régulation, de responsabilité et de confiance envers les décisions prises par les systèmes d'IA**, en s'appuyant notamment sur la possibilité d'expliquer ces décisions. Nous pouvons innover et nous approprier le créneau des technologies non investi par les acteurs des pays à faible réglementation : **explication**, validation, certification, etc.

Pour traiter ces enjeux, le groupe de travail a formulé sept recommandations principales déclinées par ordre de priorité décroissant :

1. Organiser des **challenges ouverts et des bourses aux technologies d'IA** répondant à des besoins des entreprises et de la société, **en complément des actions de fond** de montage de projets collaboratifs, de laboratoires communs et de structures agiles d'échange et de transfert pour soutenir la rencontre entre recherche et entreprises ; parmi les appels à projets, on inclura **un volet sur l'explication,** condition importante pour l'appropriation et le transfert des systèmes d'IA ;

2. Lancer un **grand projet public fédérateur en IA** comprenant notamment la réalisation d'un assistant conversationnel intelligent en soutien à la modernisation de l'action publique, mis en ligne sur les sites et les applications du service public ;
3. Soutenir la création de **plateformes communautaires, matérielles et logicielles**, dans quelques domaines prioritaires décrits ci-dessous ;
4. Confier au Laboratoire National de métrologie et d'Essais (LNE) la mission et les moyens de fédérer le **développement des méthodes de mesures, de tests et de certification pour les systèmes d'intelligence artificielle,** en partenariat avec les communautés de recherche correspondantes ;
5. **Dédier un fonds d'investissement en capital** (supérieur à 25M € par action) pour soutenir la **transformation des start-ups** dans le domaine en futurs champions ;
6. **Accélérer la mise à disposition de données publiques ouvertes pour les domaines stratégiques de la santé et de l'énergie**, afin de permettre le développement de solutions IA

qui ont particulièrement besoin d'être alimentées en données réelles pour être efficaces et pour leur adoption par les principaux acteurs économiques ;

7. **Créer une Fondation de l'IA**, dont les modalités sont à définir, afin d'incarner et fédérer dans la durée l'écosystème français de l'IA autour des défis de la société et des entreprises.

Le groupe recommande également pour la question spécifique des **plateformes communautaires**, de se focaliser sur :

- La création d'une plateforme générique **d'intégration #FranceIA**, future composante de la plateforme européenne « AI-on-demand » prévue dans le prochain appel à propositions H2020 ;
- Une plateforme **« Données, logiciels et ressources de calcul pour l'apprentissage automatique »** ;
- Une plateforme **« véhicule autonome »** réunissant les briques matérielles (y compris piste d'essais, sites, véhicules) et logicielles (algorithmes et simulateurs) permettant de développer les technologies de perception et de prise de décision pour les véhicules autonomes de demain ;
- Une plateforme de **ressources pour le traitement automatique de la langue naturelle, l'interaction humain-machine** et les agents conversationnels.

1.3 Etat des lieux et Enjeux

2.2.2 Méthodologie adoptée

Une première période d'environ un mois a été consacrée aux deux premiers points de la mission. Un élément essentiel en a été la préparation et la diffusion de questionnaires sur l'industrialisation de la recherche en intelligence artificielle. Trois questionnaires (qui figurent en annexe de ce document) ont été établis pour des populations différentes :

- Un questionnaire à destination des chercheurs, pour identifier les technologies susceptibles de faire l'objet d'un transfert à court et moyen terme ;
- Un questionnaire à destination des entreprises, pour identifier les cas d'usages industriels, les modalités à mettre en œuvre et les freins à desserrer afin de motiver et favoriser le transfert de technologies à court et moyen terme ;
- Un questionnaire à destination des intégrateurs de technologies, pour identifier les demandes et les éventuelles difficultés d'accès aux services/solutions/produits/plateformes s'appuyant sur des technologies d'intelligence artificielle susceptibles d'être proposées à des clients.

Les questionnaires ont été développés et publiés sous l'enseigne du Secrétariat général pour la modernisation de l'action publique (SGMAP) et une campagne de messagerie a été organisée afin de toucher le plus possible d'acteurs de ces différentes communautés, notamment via les têtes de réseau INRIA, CNRS, IMT, les Pôles de compétitivité, les associations savantes et professionnelles. Les formulaires ont été disponibles en ligne pendant deux à trois semaines, période courte, qui ne permet pas de réaliser une enquête en vraie grandeur, mais qui a cependant permis de collecter suffisamment de réponses pour pouvoir les considérer comme assez représentatives. Ainsi, plus de 160 réponses ont été obtenues de la part de chercheurs ; plus de cinquante entreprises ont répondu aux deuxième et troisième questionnaires.

Le principe du questionnaire est qu'il s'agit d'une auto-évaluation du potentiel – et des éventuelles difficultés – de transfert et des besoins ; il n'y a pas eu de validation croisée des déclarations faites par les répondants. Il faut donc prendre les résultats comme des indications et non comme des certitudes.

2.2.3 Identification des principaux domaines de la recherche en intelligence artificielle transférables

Vu le nombre important de réponses de chercheurs, nous avons classé et analysé celles-ci en quatre catégories, sans ordre de priorité.

a. Apprentissage, deep learning, apprentissage par renforcement

Dans la trentaine de questionnaires reçus, l'apprentissage est principalement dominé par l'apprentissage statistique, qu'il soit à base de Deep Learning ou non. D'autres approches d'apprentissage apparaissent cependant mais plus ponctuellement : traitement de l'incertitude, apprentissage développemental, apprentissage embarqué, raisonnement. Cela montre la diversité des technologies disponibles sur notre territoire.

Plus de la moitié des chercheurs ayant répondu ont des expériences de transfert technologique montrant des liens déjà forts entre académiques et industriels dans ce domaine. Sur ces transferts la majeure partie des réponses est plutôt positive ou neutre, seuls quelques questionnaires font mention d'expérience de transfert négative, en général vers des grands groupes. A ce titre, sont mis en évidence des problèmes liés à la complexité des algorithmes, leur difficulté de compréhension par les entreprises cible ainsi que le manque de confiance dans les résultats obtenus et ce, malgré les expérimentations réalisées.

La grande majorité des réponses mentionne des possibilités de transfert à court terme. Les domaines d'application sont extrêmement variés : médecine, transport, recommandation et commerce, services B2C tel que les assistants personnels ou les jeux, maintenance prédictive et problématiques industrielles, robotique/IoT et enfin tout ce qui a trait à l'embarqué.

Dans le domaine de la vision, le Deep Learning est clairement la technologie majoritaire. Là aussi les chercheurs justifient d'expériences de transfert réussies sur des applications de reconnaissance d'objets, d'animaux, etc. et aussi sur des problématiques centrées autour des activités humaines : reconnaissance de pose, reconnaissance faciale, *tracking* d'individus.

Les verrous mis en évidence pour une industrialisation des solutions développées concernent principalement l'incertitude sur la qualité des résultats, notamment sur des champs d'application nouveaux. Pour lever cette incertitude, plusieurs moyens sont identifiés. On retrouve principalement un besoin de collecter plus de données pour "augmenter les ensembles d'apprentissage" mais aussi "d'interaction avec les experts des secteurs cibles" et "d'expérimentations à grande échelle", avec un accès à des "données réelles". La capacité à "réduire l'empreinte" des algorithmes et la nécessité de pouvoir passer à l'échelle avec des architectures de calcul puissantes sont des thématiques qui reviennent plusieurs fois dans les réponses afin de résoudre plusieurs challenges actuels.

Notons également, quand cela en est fait mention sur le sujet du transfert des solutions, un réel problème de ressources pour développer une version "industrielle des algorithmes" souvent à l'état de démonstrateur ou de POC, en particulier des ingénieurs. Enfin, la problématique d'intégration des

algorithmes proposés dans des solutions plus générales (approche système) est aussi régulièrement évoquée.

b. Optimisation, planification, programmation par contraintes, aide à la décision

Dans les 40 questionnaires reçus relatifs à ce sous-domaine de l'IA, de nombreuses thématiques de recherche sont mentionnées : raisonnement en présence d'incertitude, aide à la décision, recommandation, optimisation, programmation par contraintes, recherche opérationnelle, planification, diagnostic, argumentation, jeux éducatifs, éthique et IA, robotique, identification des objets en mouvement.

Les domaines d'application considérés sont très variés. Nous pouvons citer comme exemples, la biologie, la médecine, les finances, l'intelligence économique, la vérification, la reconfiguration et la certification de logiciel, le transport, l'aéronautique, la publicité, la défense, le web, les systèmes de recommandation, le développement durable, l'agriculture, les smart-cities ou la robotique autonome.

Nous pouvons observer que certains groupes de recherche (30% des réponses) possèdent une très bonne expérience en transfert technologique, comme par exemple dans la prise de décision sous incertitude (e.g. la bibliothèque GHOST de modélisation et de résolution de problèmes d'optimisation de contraintes, développée au LS2N et utilisée par le studio de jeu français Insane Unity), la programmation par contraintes (e.g. transfert de travaux menés dans le laboratoire I3S pour la gestion automatique de musique et de textes chez Sony), ou la robotique autonome (e.g. transfert de travaux menés dans l'équipe Inria Chroma pour le véhicule autonome chez Renault et Toyota). Les industriels mentionnés sont aussi bien des grands groupes, des entreprises de taille moyenne que de start-ups. Certaines applications sont aussi réalisées en collaboration avec le secteur public.

Nous observons également que chaque groupe de recherche peut avoir des développements logiciels à des degrés de maturité variés (du démonstrateur au système testé sous conditions d'utilisation réelles). Mais dans la majorité des enquêtes, la valorisation des résultats de recherche, bien que possédant des domaines d'application prometteurs, se trouve dans une phase préliminaire. Les résultats évoqués sont de nature fondamentale mais il ne faut pas oublier que grâce à ces recherches le transfert est possible. Généralement, il y a aussi de la production de démonstrateurs ou de prototypes qui permettent de tester les concepts et les techniques induits par les résultats théoriques.

Pour aller plus loin dans le transfert et la valorisation des résultats de la recherche, les freins essentiellement évoqués sont :

- le manque de personnel dans les laboratoires pour l'implémentation, le développement et le test des résultats de recherche. La mise à disposition d'ingénieurs de recherche et de développement devrait permettre un meilleur dialogue avec les milieux industriels,

- le manque de volumes de données réelles conséquents pour la mise en place des expériences grandeur nature permettant de tester les prototypes,

- la difficulté d'acquisition des connaissances expertes métier nécessaires pour adapter des méthodes de raisonnement génériques. Sont évoqués le manque de disponibilité des interlocuteurs chez les partenaires industriels et leur trop fort turnover qui empêche le bon avancement et suivi de certains projets,

- la frilosité des partenaires industriels. Une implication et des prises de risque plus importantes de partenaires industriels devraient permettre une meilleure valorisation des résultats de la recherche. La création de laboratoires communs entre académiques et industriels, la co-conception et le co-développement de pilotes académiques-industriels sont évoqués plusieurs fois, ainsi qu'un investissement financier des partenaires industriels permettant à des chercheurs directement concernés de travailler sur le transfert (l'encadrement de doctorants sur un contrat de thèse en milieu industriel ne pouvant avoir en général que des résultats limités).

c. Connaissances, raisonnement, ontologies, web sémantique, TALN, émotions, chatbots

La quarantaine de réponses au questionnaire relevant de cette catégorie mentionne plusieurs des thématiques communes avec les deux précédentes telles que l'apprentissage automatique ou le raisonnement dans l'incertain, mais font surtout ressortir comme thème prédominant le sujet du traitement automatique du langage naturel (y compris vocal), tant pour la fouille de données textuelles que pour l'interaction humain-machine. D'autres sujets émergent également, tels que la représentation et l'ingénierie des connaissances, le calcul de préférences et de recommandations, les agents virtuels, le calcul émotionnel et le web sémantique.

Concernant les domaines d'applications cible, la plupart des secteurs requérant de la relation humain-service/système sont visés : commerce électronique, ciblage et recommandation, e-réputation, formation et e-éducation, e-santé, assistants personnels, agents conversationnels, robots compagnons et industriels, traduction et résumé automatiques, suivi de réunions, etc.

Dans cette catégorie, peu d'équipes de recherche font état d'expériences de transfert, réussies ou non. Par ailleurs, très peu des systèmes mentionnés sont indiqués comme ayant été testés en condition d'utilisation réelle. Ce qui, par exemple, dans le domaine du traitement automatique du langage naturel, pose question au vu des besoins exprimés par les entreprises, d'autant qu'il s'agit d'un thème de recherche historique et pour lequel certaines applications opérationnelles sont d'ores et déjà déployées.

Quant aux moyens requis pour l'industrialisation, le renforcement des ressources d'aide au développement et aux expérimentations en vraie grandeur (ingénieurs, plateformes, données réelles) est particulièrement évoqué, en particulier à travers des programmes de soutien à la maturation et au transfert. Une autre préconisation mérite d'être mentionnée : réduire le nombre d'intermédiaires dans les relations Recherche-Industrie en favorisant l'accueil d'ingénieurs de grandes entreprises ou de PME dans les laboratoires de recherche et, inversement, le détachement de chercheurs dans l'industrie pour accompagner les transferts (« avec des rémunérations libres », sic).

d. Evolution artificielle, systèmes multi-agents, robotique, autres

Cette section regroupe des sous-domaines qui partagent notamment une vision distribuée de l'intelligence artificielle, que ce soit sous forme de populations de solutions qui évoluent pour apporter la meilleure réponse possible à un problème d'optimisation ; de réseaux d'agents intelligents collaborant pour simuler des systèmes complexes ou pour combiner des solutions partielles ; de systèmes ouverts intégrant des agents intelligents coopérant entre eux et avec des humains pour fournir des services à valeur ajoutée ; ou de robots interagissant entre eux ou avec des humains.

Plus de quarante technologies ont été identifiées par les répondants au questionnaire dans cette catégorie.

Une dizaine de technologies utilisant **l'évolution artificielle (EA)** a été mentionnée ; le plus souvent, il s'agit de démonstrateurs. Les applications sont variées, l'EA étant connue pour pouvoir résoudre des catégories de problèmes très diverses. On peut noter en particulier des applications en planification et ordonnancement, en reconnaissance de mouvements, en robotique, ou pour les jeux vidéo ; notons également l'application de l'EA pour détecter des bugs dans des logiciels, ce qui pourrait offrir un marché intéressant à une société qui voudrait se développer sur cette base. Dans la plupart des cas, les besoins identifiés sont de travailler avec des utilisateurs pour transposer les technologies à des cas réels et développer les codes pour pouvoir passer à l'échelle.

Une quinzaine de contributions relève des **systèmes multi-agents (SMA).** Cette technologie est principalement utilisée pour trois familles d'applications : la prise de décision décentralisée, la simulation de systèmes complexes avec de nombreux composants/comportements en interaction, et plus largement l'intégration et la coopération de systèmes intelligents dans un cadre ouvert et décentralisé, par exemple pour la négociation entre entités antagonistes. Ces trois familles s'illustrent dans de multiples domaines tels que l'Internet des objets, la santé, les jeux, la gestion intelligente de l'énergie, la production flexible dans l'industrie du futur, les villes intelligentes, le transport, etc.

D'autres utilisations sont également mentionnées comme la population de mondes virtuels pour les jeux vidéo ou la gestion distribuée d'ontologies en aide à la conception. Plusieurs des systèmes sont déclarés comme testés en conditions réelles, et demandent essentiellement des travaux d'ingénierie logicielle pour rendre le système plus robuste et passer à l'échelle sur un plus grand nombre de cas.

Le monde des SMA est aussi caractérisé par les *plates-formes multi-agents*, des environnements logiciels qui permettent la conception, la programmation et la mise en œuvre de SMA pour des applications. Plusieurs plateformes existent sur le territoire national ; il y a là un potentiel d'industrialisation intéressant qui pourrait aussi fédérer un ensemble de technologies au service d'une communauté et plus largement pour l'intégration des technologies développées dans les autres sous-domaines de l'IA. Enfin, à noter une utilisation des SMA dans le domaine de la perception d'environnement complexe pour un véhicule autonome : en plus des demandes classiques sur la fiabilisation du système, il y a également des demandes de standardisation et d'évolution de la réglementation pour permettre des expérimentations.

Quelques contributions relèvent de la **robotique**, mais on peut affirmer que le nombre de réponses reçues est très inférieur à ce qu'il aurait pu être compte tenu du grand nombre d'acteurs, chercheurs et industriels, actifs sur ce sujet. Mentionnons pour mémoire des travaux sur la localisation, sur le placement de robots, sur la planification de mouvements. La faible représentativité des réponses ne nous autorise pas à conclure quoi que ce soit en la matière.

Enfin, parmi les autres technologies appartenant à cette catégorie, nous pouvons relever des sujets très divers, souvent motivés par des considérations éthiques et morales comme la supervision de trading haute fréquence par des « gendarmes financiers numériques » ou une plateforme d'intelligence collective pour la supervision de débats ; diverses applications de fouille de données (biologie, chimie, musique), des systèmes intelligents distribués pour améliorer la performance des systèmes d'information de type ERP, ou encore des environnements virtuels adaptatifs basés sur la compréhension des intentions de l'utilisateur. La plupart sont à l'état de démonstrateur ou preuve de concept, et demandent donc des développements significatifs pour se rapprocher d'une utilisation industrielle à moyen terme.

2.2.4 Le point de vue des entreprises

Les questionnaires destinés aux entreprises et intégrateurs ont reçu une cinquantaine de réponses, souvent assez précises et pour quelques-unes, très détaillées. Ils ont été diffusés au travers du MEDEF et du CIGREF, vers les industriels en lien avec les associations savantes de l'intelligence artificielle et de la recherche opérationnelle, ainsi que vers les sites *facebook* de l'écosystème des start-ups françaises et via le SYNTEC Numérique et Tech In France à l'ensemble des ESN, sociétés de services et de conseil en technologie et éditeurs de logiciels.

a. Penser chacun des maillons reliant la recherche, l'offre et la demande

Le questionnaire destiné aux intégrateurs, conçu comme complémentaire à celui des entreprises utilisatrices d'IA, nous aide à interpréter le côté « demande » de la filière. Il apparaît en effet nécessaire d'articuler consciemment l'adéquation *push/pull* de l'offre, issue de la recherche transférée aux start-ups, avec la demande de nos industries et sociétés de services afin d'en garantir la mise en œuvre opérationnelle. Les sociétés de conseil et les intégrateurs pourraient être des acteurs efficaces du couplage de l'offre et de la demande, contribuant à l'efficacité et à la cohérence d'une filière intégrée. Grâce à leur maillage du territoire et de l'écosystème, ces sociétés peuvent contribuer à l'évangélisation/éducation des entreprises et être un relai efficace des start-ups qui élaborent les offres produit et service au contact des laboratoires de recherche. L'agilité de la filière ressort comme un élément clef de la performance de cette diffusion des technologies.

Différences d'éducations et de densités de connexion
L'écosystème national des start-ups bénéficie d'une homogénéité culturelle et d'une densité de connexion qui favorise et facilite l'utilisation en B2B de l'offre IA issue de nos start-ups technologiques, par les start-ups sectorielles qui en auraient l'usage. Le tissu des grands groupes, des ETI et des PME, par contre, ne bénéficie pas de cette immersion naturelle dans la technologie issue des laboratoires et est donc plus soumis à la pression marketing, publicitaire et commerciale des acteurs nord-américains. Eduquer et développer la demande, tout en l'orientant vers les solutions issues de la recherche nationale nous semble être un élément important du développement de la filière. La population des directeurs de l'innovation et plus largement des directions générales pourrait par exemple bénéficier d'une éducation/évangélisation plus poussée qui lui donnerait les moyens de connaître, branche par branche, les applications utiles, d'en comprendre les niveaux de risque, et d'identifier les acteurs – en particulier nationaux – à même de les servir au meilleur rapport qualité-prix.
Faute d'une attention particulière à ces maillons, on peut penser que lorsqu'elle cristallisera, la demande s'orientera naturellement vers les solutions nord-américaines dont les machines de marketing, communication et vente dominent actuellement le paysage.

Faire des exigences sociales un atout ?
L'exigence européenne et en particulier française sur la protection des données et de la vie privée sont le plus souvent présentées – ou vécues – par les entreprises comme un frein à leur développement, qui les empêche de proposer certains produits innovants et qui les pénalise vis-à-vis de leurs marchés potentiels. Il peut en être de même pour des exigences de validation garantissant la qualité des algorithmes, de méthode de certifications, ou d'explicabilité des résultats. Ces spécificités nationales pourraient devenir des atouts si l'état accompagne sa fermeté réglementaire par des dispositifs aboutissant à la création et à la commercialisation des solutions innovantes permettant aux industries et services de respecter simplement ces règlements. La réglementation se globalisant, les acteurs nationaux de cette niche « explication, validation, certification » seraient prêts à commercialiser au plan international des solutions adéquates et validées.

b. Les domaines d'application

On détecte un certain nombre de régularités tant sur les domaines d'applications que sur les méthodes et leurs modalités de mises en œuvre. On trouve ci-dessous une première liste issue du sondage effectué auprès des industries et services.

- **Voiture autonome et connectée** : les deux sujets sont distingués car ils mobilisent des technologies différentes et ciblent des objectifs distincts : automatiser tout ou partie de la conduite d'un véhicule dans un cas et fournir des services connectés en mobilité dans l'autre.
- **Contrôle qualité, conformité et traçabilité** : ces applications permettent de baisser considérablement le coût de la mise en œuvre de normes qualité. Différentes techniques peuvent être implémentées, comme celles des règles métiers pour assurer la traçabilité ou la

conformité, ou encore des techniques d'interprétation d'images ou de signaux pour le contrôle de la qualité d'objets manufacturés.

- **Interprétation temps réel de situations complexes** dans un cadre d'aide à la décision : ces applications complexes peuvent combiner plusieurs technologies d'IA, comme la vision, la sémantique, le raisonnement, et être intégrées par des systèmes multi-agents.
- **Relation client** : le domaine est vaste et va de la poursuite de l'automatisation des tâches de routine du CRM démarrée depuis une dizaine d'années – application automatisées de programmes de promotions, création de lettres personnalisées, *phoning* automatisé, classification des clients – aux *chatbots* apprenants qui automatisent les interactions de conseil, support et vente via le web et les mobiles.
- **Langage et interaction** : cette dimension est transversale, elle est également centrale car elle nécessite le meilleur niveau d'automatisation du traitement du français notamment et des interactions dans les cadres et les codes culturels français.
- **Agents conversationnels** (ou *Chatbots*): ils sont utilisés pour simuler un dialogue naturel entre un service automatisé et un humain. Le service peut être incarné dans un robot, ou virtuel, en ligne sur le web ou sur un smartphone. Les agents conversationnels interprètent et produisent de la parole, du texte, ou des langages multimodaux. Ils doivent reconnaître les états émotionnels, les intentions des utilisateurs et qualifier leur niveau de compétence ; ils peuvent également avoir à interpréter leur langage corporel. Selon les cas, ces agents conversationnels peuvent être au service de l'entreprise automatisant la relation avec le client ou au service de la personne, en tant qu'assistants personnels. On peut les trouver également en support aux réunions virtuelles.
- **Conception/configuration** : il s'agit de concevoir un artefact de façon automatisée ou semi-automatisée et en optimisant certains de ses aspects, par exemple son coût. L'artefact est le plus souvent conçu à partir d'une bibliothèque de composants pouvant contenir de quelques dizaines à quelques milliers d'entités. Il peut être virtuel, comme un contrat ou l'organisation d'un voyage, ou matériel comme une voiture ou un ordinateur.
- **Logistique** : il s'agit ici d'optimiser en différé ou en temps réel l'ordonnancement d'ateliers de production, la gestion et l'approvisionnement d'entrepôts ou de magasins, ou encore de planifier des activités complexes telles qu'un chantier de bâtiment ou la construction d'un avion.
- **Maintenance** : prédictive et corrective. On cherche ici à comprendre et corriger les pannes, ou de préférence à les anticiper grâce à l'utilisation de capteurs dont les informations sont comparées à des modèles prédictifs produits lors de la conception de l'artefact ou créés par apprentissage.
- **Interprétation d'événements** : c'est le cas de l'interprétation de situations évoluant dans le temps ; on cherche à détecter des évènements signifiants ou à interpréter des enchaînements porteurs de sens. Les deux grands champs d'exploitation sont les opérations militaires et celles des marchés financiers comme le *trading* à haute fréquence - celle des IA déployées sur des supercalculateurs ou la spéculation sur les matières premières. On met en œuvre un mélange d'apprentissage, de modèles mathématiques et de règles « métiers », qui permettent la reconnaissance de séquences d'évènements.
- **Arts** : l'art n'est pas en reste avec l'aide à la sélection et l'enchaînement, ou à la composition et l'interprétation de musique. Les textiles, les arts graphiques, la photographie et la vidéo sont également des champs d'applications croissants de l'IA. On peut faire une mention particulière pour le domaine de la modélisation et de l'animation 3D, un domaine sur lequel la créativité française est très appréciée et dont les liens avec l'IA qui s'intensifient sont déjà incarnés dans certaines start-ups.
- **Calcul de Prix :** prévision et fixation des prix à partir de modèles prédictifs mathématiques ou appris, et de synthèses d'informations en temps réel. Une application inventée pour la fixation des prix du transport aérien s'est progressivement généralisée à de nombreuses industries liées à la mobilité : trains, VTC, hôtels, loisirs. On la trouve également déployée chez les start-ups qui servent les nouveaux marchés de la location et de l'occasion.

- **Médecine** : le domaine est vaste ; on peut distinguer d'un côté l'interprétation de signaux radiographiques ou électro-physiologiques, la segmentation d'images, et de l'autre le télé-*monitoring* qui mobilise des méthodes d'interprétation de signaux issus de capteurs temps réel afin d'interpréter à distance le comportement et l'état d'une personne supervisée, voire de poser un diagnostic. Plus généralement, l'IA va jouer un rôle important pour la médecine personnalisée, notamment par les méthodes que sous-tend la bioinformatique.
- **Gestion des connaissances** : c'est le domaine de prédilection des méthodes développées dans le cadre du web sémantique, utilisées ici pour représenter des connaissances ou des expertises à sauvegarder, partager, évaluer et parfois automatiser.

c. Technologies mobilisées

Au sein des technologies mobilisées, **l'apprentissage** se taille une belle part, et en particulier le *deep learning*, mais on trouve également les technologies de compréhension et de production du langage, notamment du **Français,** et de conduite du **Dialogue Naturel**, les **Ontologies,** les technologies du **Big Data**, l'**Optimisation** et la **Programmation par Contraintes**, l'**Analyse d'Images** et des **flux vidéos**, ainsi que le **Calcul Emotionnel** et les **Règles Métiers**, ou encore les technologies d'IA banalisées depuis une dizaine d'années pour l'automatisation des décisions « de routine ».

d. Mise en œuvre, modalités et freins

Les stades de mises en œuvre sont assez variés. On trouve des applications déployées en exploitation opérationnelle quotidienne comme par exemple le calcul de prix, mais il apparait qu'un grand nombre d'entre elles en sont encore au niveau de l'expérimentation. Les entreprises semblent être au stade de la découverte de l'IA pour une large part d'entre elles ; elles éprouvent des difficultés à aborder la diversité et la complexité des méthodes et semblent manquer de poissons pilotes pour les aider à naviguer entre les applications faciles et les expérimentations plus audacieuses. Cet état de faible compétence des décideurs peut les exposer au risque de *staffing* inadéquat des équipes internes ou à adhérer à des discours de communication faisant miroiter des solutions dont les bénéfices pourraient être décevants. L'IA ayant déjà connu un hiver industriel, il est important de prendre garde à d'éventuelles désillusions.
Les modalités de mise en œuvre quant à elles recouvrent un spectre très large, allant du développement en interne à l'appel à des sociétés de conseils spécialisées ou à des intégrateurs, en passant par des collaborations régulières avec des laboratoires de recherche, parfois soutenues par des bourses Cifre ou des projets collaboratifs subventionnés. Nous ne notons pas de spécificité IA par rapport au problème général d'adoption de nouvelles technologies.
Quant aux freins mentionnés, au-delà des classiques difficultés de **financement**, on trouve l'**indisponibilité d'ensemble de données**, fiables, adaptées et prêtes à l'emploi et du **savoir-faire** pour en tirer parti, la **crainte et la peur** de certains utilisateurs, et le manque **de confiance dans les résultats** des méthodes d'IA, en particulier en ce qui concerne la **justesse**, la **robustesse** et la **validité** des résultats. Sans surprise, l'**adaptation aux réglementations** est perçue comme un handicap typiquement français, ainsi que la difficulté d'accéder à des **compétences opérationnelles**, et l'**absence d'infrastructures** et de moyens de **calcul faciles d'accès et d'usage**.

2.2.5 Le point de vue des intégrateurs, sociétés de services et de conseil

Les intégrateurs, sociétés de services et de conseil sont à l'interface des besoins clients et des technologies. Leur rôle d'intermédiaires entre, d'une part, les chercheurs et start-ups développant des technologies et, d'autre part, les clients utilisateurs, est très important dans la chaîne de valeur de l'intelligence artificielle (IA), et recouvre plusieurs aspects :

- Ils assurent la convergence entre les besoins et les technologies développées en intervenant auprès des concepteurs en amont pour orienter ou susciter des productions cohérentes avec les attentes du marché, et en aval pour structurer la demande.
- Leurs interventions orientées métier leur permettent d'identifier des champs d'application d'une même technologie dans des contextes différents.
- Ils adressent un marché large (collectivités, PME, ETI, grands groupes) où tous les acteurs ne sont pas forcément sensibilisés aux usages possibles de l'IA, d'où des interventions en amont des projets opérationnels sur des cursus de formation, ou dans le montage de projets de recherche (ANR, CIFRE, H2020, etc.).
- Ils mettent en œuvre des solutions associant des technologies complémentaires de l'IA associées à d'autres technologies ou infrastructures.
- Ils participent à la maturation des technologies en transformant les algorithmes issus de la recherche en solutions performantes et intégrables dans les environnements clients.

Le rôle des intégrateurs est ainsi clef dans la diffusion et l'appropriation des technologies autour de l'IA, ainsi que dans l'adéquation des solutions aux besoins des clients.

Les intégrateurs font face à des problématiques d'interface sur deux niveaux :

- Premièrement, et au même titre que les grands groupes, certains intégrateurs possèdent une faible visibilité sur les solutions IA existantes qu'il est possible de combiner. Ils privilégient ainsi les solutions considérées comme matures et commercialisées par les GAFA ou IBM, au risque parfois de laisser de côté des solutions françaises ou européennes plus adaptées.

- Deuxièmement, les intégrateurs possèdent rarement une interface identifiée ou efficace avec les différents domaines de la recherche, d'où un manque d'efficacité sur deux volets : certains besoins clients mal identifiés par les laboratoires de recherche, et des travaux de recherche pas toujours valorisés au niveau des entreprises. Des interactions clients-chercheurs existent dans des domaines de pointe comme l'industrie de la défense, mais peu d'interfaces efficaces existent dans le domaine des services, ou celui des banques/assurances, du commerce, ou encore au niveau du tissu des PME/ETI.

a. Axes d'amélioration

Au même titre que la filière IA, l'écosystème des intégrateurs et des sociétés de services et de conseil n'est pas monolithique, mais composé d'une variété d'acteurs.

Les segmentations significatives de cet écosystème par rapport à l'identification des freins et des besoins sont :

- Le métier client ciblé par l'intégrateur : les relations entre entreprises, intégrateurs, start-ups et chercheurs sont différentes dans l'industrie, le commerce, la relation client, les banques/assurances, ou la santé.
- La taille de la structure : les structures importantes sont souvent plus à même de financer des grands projets et d'investir sur des programmes pluriannuels, contrairement aux petites structures qui peinent à bénéficier des différents financements de projets nationaux ou européens, et à créer des liens solides avec le monde de la recherche.
- La maturité des technologies utilisées par l'intégrateur : un intégrateur sera plus autonome sur une technologie mature. En revanche, les technologies émergentes nécessiteront beaucoup plus d'interaction avec la recherche ou les start-ups.

Cependant, certains invariants se dégagent pour l'ensemble des catégories consultées : les besoins critiques identifiés dans le cadre de discussions approfondies avec certains intégrateurs et issus de

l'analyse des réponses aux questionnaires mettent en évidence trois axes d'amélioration du transfert de technologie des laboratoires vers les acteurs économiques via les intégrateurs, quel que soit le secteur, la taille ou bien les technologies manipulées par ces derniers :

1. **Des problématiques d'agilité de la filière** :
   a. Un manque de visibilité des formations, des acteurs, des compétences, des technologies et de leur niveau de maturité.
   b. Des difficultés à activer des cycles courts de collaboration (entre 3 et 9 mois) et d'accéder lors de ces cycles à de l'expertise (ingénieurs-chercheurs des instituts, post-doctorants, etc.)
   c. Un manque de visibilité sur la manière de faire émerger des collaborations et converger les efforts des différents acteurs de la chaîne de valeur de l'IA.
2. **Des problématiques autour des API/processus/données** :
   a. La question de se lancer ou non dans la création de plateformes concurrentes à IBM, Microsoft, etc. de modules technologiques et d'hébergement et analyse des données.
   b. L'urbanisation des composants d'une éventuelle plateforme.
   c. La régulation de l'hébergement des données et de leur sécurisation en phase avec les valeurs européennes.
3. **L'évangélisation de l'écosystème et des utilisateurs** :
   a. Le manque de communication structurée auprès de certaines catégories de clients potentiels (milieu agricole, médical, sportif, etc.).
   b. Les messages fortement anxiogènes sur l'impact du développement des IA sur les individus et les salariés.
   c. Les risques d'écart entre les attentes suscitées auprès des clients et les résultats effectifs à court terme.

b. Propositions d'actions sur les différents axes

- **L'agilité de la filière** : visant une dynamique de la demande en convergence avec l'offre technologique, et en plus d'une nécessaire cartographie des acteurs, la capacité à mobiliser les expertises sur des cycles courts est clef dans le processus de transfert. La définition d'un processus allégé pour les financements de quelques centaines de milliers d'€ mobilisables rapidement sur des projets courts (entre 3 et 6 mois) serait de nature à dynamiser la filière.

- **L'accès aux Données et Plateformes** : les outils et données aujourd'hui disponibles sont issus des GAFIM[1] et n'apportent pas ni le niveau de performance, ni l'explicabilité attendus par le marché européen ; au-delà des enjeux de souveraineté, la structuration d'environnements permettant cette accessibilité est critique pour répondre aux attentes des entreprises avec les technologies locales. L'incitation de l'écosystème à déployer des projets communs client-intégrateur de création et de documentation d'API avec ouverture des données non-sensibles du client aux autres acteurs du marché pourrait permettre le déploiement de ces technologies

- **L'évangélisation des acteurs** : d'un côté, les attentes sidérantes évoqués par la presse et la science-fiction surfent sur les aspects anxiogènes développés tant sur les dimensions du travail que de la révolution de la société jusqu'à la disparition de l'humain. D'autre part, le réalisme

[1] Google, Amazon, Facebook, IBM et Microsoft

pragmatique de ce qui est aujourd'hui effectivement opérationnel semble prometteur mais peu avenant. Entre les deux, un juste milieu est nécessaire pour à la fois susciter une demande « raisonnée » de ces technologies et éviter un rejet de principe sur des outils qui autorisent un développement industriel significatif.

Par ailleurs, les grands intégrateurs doivent avoir un rôle clef dans la définition et la mise en œuvre de la filière d'IA française et la mise en place des outils et des processus permettant sa performance et sa compétitivité.

1.4 Recommandations pour stimuler et soutenir les coopérations entre entreprises et laboratoires favorisant le transfert

Un programme national ambitieux sur l'intelligence artificielle

La France doit saisir au mieux les bénéfices de l'intelligence artificielle afin d'assurer la croissance et la compétitivité de son économie, favoriser par le numérique la transition énergétique et réduire, pour des raisons évidentes de souveraineté et de sécurité nationale, sa dépendance aux solutions et technologies proposées par des acteurs étrangers.
Pour ce faire, nous proposons le lancement d'un programme national ambitieux sur une période de dix ans qui viserait à :

1) Stimuler la recherche nationale, favoriser la maturation des technologies et fédérer les acteurs de la filière au sein d'écosystèmes entre recherche académique, technologique et entreprises ;
2) Assurer en amont la confrontation entre la recherche et les cas d'usages des entreprises, pour permettre la meilleure prise en compte possible de leurs besoins et contraintes ;
3) Favoriser la diffusion et le transfert des résultats et technologies issus de la recherche dans les entreprises, PME, ETI et grands groupes ;
4) Favoriser la création de start-ups dans les domaines spécifiques de l'IA et accompagner leur croissance indispensable à l'émergence de « champions de demain » ;
5) Mettre en place des plateformes communautaires permettant notamment le parangonnage et l'expérimentation ;
6) Mettre en place des plateformes d'intégration et de démonstration des innovations des technologies IA ;
7) Développer des programmes de formation initiale et permanente, y compris chaires académiques et industrielles ;
8) Mettre en place des dispositifs à l'échelle nationale pour fédérer les communautés afin de participer aux programmes européens, H2020 et sa suite ;
9) Accompagner ces réalisations d'une réflexion sur l'impact sociétal de l'IA et d'un dialogue citoyen en relation avec les autorités de médiation et de régulation.

Le programme comportera de plus un volet de communication/médiation afin d'assurer la diffusion de ses résultats. La création d'un institut de l'intelligence artificielle est également proposée.

Les financements proviendraient à 25 à 35% des entreprises et collectivités territoriales, mobilisant des fonds privés et publics, s'appuyant notamment sur les programmes existants en particulier le PIA3.

Le programme serait doté d'une gouvernance appropriée, faisant appel à des représentants de l'Etat, des opérateurs de recherche et des industriels.

Un volet relatif à la valorisation et aux transferts des résultats et technologies issus de la recherche

Nous nous focalisons dans ce document sur ce volet du programme sachant que d'autres groupes traitent les aspects complémentaires (recherche fondamentale, formation, dialogue citoyen en particulier).
Nous proposons sept actions principales, certaines étant composées de sous-actions.

Recommandation 1 : Transfert recherche-entreprises

- **Organiser des challenges ouverts et des bourses aux technologies d'IA** permettant à une entreprise, en particulier les PME/ETI, voire à l'administration, de lancer un challenge auprès de start-ups ou solliciter le soutien d'organismes de recherche. L'objectif serait de répondre à une difficulté technologique ou évaluer l'intérêt potentiel (y compris d'usage) d'une technologie avant son implémentation.
- **appels à projets « multipartenaires »** impliquant obligatoirement a minima un organisme de recherche et un industriel. Sélection par l'excellence et l'impact économique de projets de recherche industrielle utilisant des technologies existantes dans les laboratoires de recherche, pour les faire passer à la dimension industrielle, soit un passage de TRL 3 à 6 (i.e. *Technology Readiness Level,* niveau de maturité technologique*)*.
  Une condition nécessaire serait la démonstration rapide d'une plus-value mesurable pour un des secteurs utilisateurs listés ci-dessus, et la participation effective d'entreprises du secteur y compris au niveau du pilotage des projets.
- **Appels à projets de type « laboratoire commun »** entre académique et entreprise dont l'objectif serait là aussi l'industrialisation de résultats de recherche ou de technologies.
- **Appels à projets sur l'explication** ciblés sur les besoins des entreprises et des administrations.

Recommandation 2 : MarIAnne, Grand projet public IA fédérateur

- Ce grand projet vise à développer notamment un assistant conversationnel intelligent en soutien à la modernisation de l'action publique, mis en ligne sur les sites de service public, utilisant les ressources en ligne et les nombreuses API réalisées par la Direction interministérielle du numérique et du système d'information et de communication de l'État (DINSIC).
- Cette initiative permettra à nos services publics de ne pas dépendre des outils fournis par des acteurs étrangers pour l'exploitation de nos données publiques et personnelles gérées par les administrations, d'éviter la perte d'une part croissante de la valeur ajoutée à leurs profits et d'offrir une solution d'intégration des briques technologiques développées en France.

Recommandation 3 : Plateformes communautaires
Appels à projets pour des plateformes logicielles et/ou matérielles visant notamment

- **(1)** au parangonnage et à l'expérimentation voire à l'organisation de challenges ou à l'intégration et à la démonstration des innovations. Ces plateformes auront principalement trois fonctions :
    - (i) fournir des données réalistes et représentative des problèmes et usages réels, en quantité et en qualité, ainsi que des capacités d'exploitation et d'expérimentation – par exemple des sites de type « living labs » – permettant de tester des nouveaux services, des robots domestiques, des véhicules autonomes, etc. ;
    - (ii) fournir des ressources logicielles et matérielles permettant notamment le passage à l'échelle des méthodes et algorithmes sur les données mises à disposition;
    - (iii) être équipées de capteurs, d'actionneurs, et d'analyseurs permettant de mesurer les résultats des expérimentations et ainsi de parangonner les technologies.
- **(2)** au développement de chaînes d'outils et de méthodes pour la conception, l'intégration ou la validation de briques logicielles et de systèmes complets

Des plateformes identifiées sont proposées en raison de leur importance particulière et du potentiel de transfert qu'elles peuvent soutenir (voir section 2.3 pour plus de détail):

- Plateforme d'intégration et de démonstration des innovations, composante de la plateforme européenne AI-on-demand ;
- Plateforme de données, logiciels, ressources de calcul pour l'apprentissage ;
- Plateforme véhicule autonome (en relation avec le travail du GT dédié) ;
- Plateforme de ressources pour le traitement automatique de la langue naturelle, l'interaction homme-machine et les agents conversationnels.

Recommandation 4 : développement de méthodologies de tests et de certification
Un enjeu fondamental pour le développement de l'IA en France est de mettre en place l'ensemble de la chaîne d'outils (protocoles, méthodes, simulateurs, environnements de conception et d'intégration, etc.) permettant d'expliquer, de garantir et de certifier les technologies utilisant de l'IA. Il est proposé que cette action soit mise sous la responsabilité d'animation du Laboratoire National de métrologie et d'Essais (LNE) dont le domaine de l'intelligence artificielle est un des axes stratégiques de développement. Voir détail en section 2.2.

Recommandation 5 : Dédier un fonds d'investissement en capital
D'un montant supérieur à 25M€ par action, ce fonds est destiné à soutenir la **transformation des start-ups** dans le domaine en futurs champions.

Recommandation 6 : Initiatives spécifiques aux secteurs santé et énergie ;

- Santé : soutenir la **constitution de bases de données médicales et de santé**, en étant attentif aux questions réglementaires et de vie privée ; mettre en place un fonds d'investissement dédié aux technologies d'IA pour la santé ;
- Energie : développement d'un **avatar temps-réel du réseau** s'appuyant sur **une matrice de visibilité** de l'ensemble des données, afin de soutenir en particulier la mise en place de deux grands défis : le développement d'une **solution d'équilibrage**, et la conception de **places de marché** pour l'énergie.

Recommandation 7 : Créer la Fondation de l'IA

- Dans des modalités restant à définir, la Fondation vise à incarner et fédérer dans la durée l'écosystème français de l'IA autour des défis de la société et des entreprises.
  Elle sera un lieu privilégié d'échanges et de diffusion sur les avancées, potentialités, opportunités et risques liés à l'intelligence artificielle. Elle permettra aux chercheurs, au grand public et aux industriels de mieux appréhender les grandes questions posées par l'IA.

2. Approfondissements thématiques

2.1 Approfondissements sur les thèmes du test, de l'expérimentation et des plateformes

2.1.1 Propositions de développement de tests et expérimentations

Un enjeu fondamental pour le développement de l'IA en France est de mettre en place l'ensemble de la chaîne d'outils (protocoles, méthodes, simulateurs, environnements de conception et d'intégration, etc.) permettant d'expliquer, de garantir et de certifier les technologies utilisant de l'IA.

L'estimation de la qualité des technologies de l'IA s'effectue à l'aide de protocoles de mesure qui varient en fonction de la maturité même des technologies. Ces dernières résident toutefois sur des fondements communs qui sont ceux de la métrologie à savoir : reproductibilité, répétabilité, estimation de la justesse, sachant que la recherche de preuve n'est pas obligatoire. Il s'agit de comparer des approches dans les mêmes conditions sur une tâche précise via des tests et des expérimentations. Notons qu'elles requièrent le développement de nouvelles technologies de métrologie, de validation ou des simulateurs adaptées à ces systèmes spécifiques.

Enfin, la diffusion de ces technologies aux industriels s'accompagne obligatoirement de développements de nouvelles méthodes et de nouveaux outils pour la conception et l'intégration spécifiques aux systèmes à base d'IA, y compris en considérant les architectures matérielles qui supportent les algorithmes.

2.1.1.1 Objectifs des tests et expérimentations

Les tests et expérimentations cherchent à répondre aux objectifs suivants :

- **Assurer le transfert aux entreprises**

Il s'agit là de permettre aux industriels de connaître les capacités exactes des systèmes d'IA à répondre à la tâche attendue et ainsi de leur donner confiance dans la pertinence des réponses fournies par ces systèmes. Les tests et expérimentations doivent permettre de connaître l'état de l'art des systèmes en comparant les solutions entre elles ou à une référence. Ils doivent également permettre d'évaluer les progrès des systèmes au cours du temps en identifiant les facteurs d'influence.

- **Fournir des référentiels communs aux développeurs**

Pouvoir se comparer dans le même cadre expérimental (données, plateformes, métriques…) permet aux développeurs de se situer vis-à-vis de l'état de l'art. Ce cadre expérimental est soumis à plusieurs contraintes. Il doit être facilement accessibles, respecter la législation du droit des données et des personnes (droit d'auteur, anonymisation, …), être représentatif des cas d'usage d'autant plus que les technologies mises en jeu sont matures.

- **Créer un climat de confiance**

La mise en œuvre des tests et expérimentation est une tâche à part entière qui devrait être confiée à un organisme tiers indépendant en particulier pour les technologies peu matures où la recherche publique est particulièrement présente. Les tests et expérimentations peuvent également être mis en œuvre par les industriels autour de leurs enjeux propres lorsque les technologies sont plus matures. La définition des tests doit se faire en collaboration avec l'ensemble de la communauté (chercheurs, industriels, tutelles, utilisateurs finaux, etc.), Elles doivent s'inscrire dans une démarche de développement d'une certification propre aux systèmes à base d'IA. Cette certification concernera non seulement les aspects organisationnels des développements mais également le niveau de fiabilité des systèmes.

- **Développer et fournir aux entreprises les technologies et les outils indispensables pour expliquer, garantir et certifier**

Il s'agit de soutenir le développement de nouvelles méthodes et de nouveaux outils autant pour répondre aux enjeux de la « métrologie », de simulation ou de la validation des systèmes à base d'Intelligence artificielle, que pour offrir, notamment aux entreprises, des environnements de conception et d'intégration, y compris sur les architectures matérielles, en vue d'améliorer la performance et la fiabilité des systèmes. Est également essentielle la simulation des comportements des systèmes en conditions réelles dans des installations spécifiques. Notons que ces activités pourront conduire à la création d'un écosystème et donc à produire sur le territoire national de la valeur ajouté, y compris en terme d'emplois.

2.1.1.2 Recherche et développement d'une métrologie propre aux systèmes à base d'IA : obtenir les meilleurs méthodes et critères d'évaluation

Pour pouvoir qualifier correctement les systèmes d'IA, il est nécessaire de doter la France de méthodes communes de mesures. Ces méthodes passent par le développement de plateformes d'évaluation communes permettant d'assurer la reproductibilité, la répétabilité et estimer la justesse des réponses des systèmes. Les modes d'évaluation peuvent prendre des formes différentes en fonction des cas d'usage à traiter. L'intervention humaine dans le cadre de l'évaluation est une thématique qu'il est nécessaire de mieux définir. Certaines technologies peuvent être évaluées à l'aide de références physiques, d'autres par des annotations humaines ou encore à l'aide de simulations. Mieux structurer ces évaluations et expliciter dans quel cadre utiliser ces différentes techniques est nécessaire pour bien qualifier les technologies. L'estimation de la significativité des différences de qualité, l'interprétabilité des métriques, mieux expliquer les prises de décisions restent des problèmes de R&D à résoudre. Le développement d'une équipe française ayant comme responsabilité de proposer les meilleurs critères d'évaluation, comme le fait le NIST aux Etats-Unis, permettrait de mieux structurer la filière de l'IA française et de mieux se repérer dans les offres tout en protégeant la souveraineté nationale.

2.1.1.3 L'organisation de challenges : permettre de se comparer en toute confiance

Les challenges sont des projets pluriannuels qui consistent à proposer un cadre commun de comparaison des solutions de systèmes. Cette organisation existe déjà dans certains domaines (recherche opérationnelle[2], traduction[3], traitement de documents scannés[4], repérage de personnes[5], négociation[6], réputation et confiance[7], robotique[8], systèmes multi-agents[9], véhicule autonome[10], …). Il est nécessaire de la généralisation à tous les domaines de l'IA en confiant cette mission à un organisme public garant de l'équité entre l'ensemble des participants : l'évaluateur.

Les challenges consistent à comparer les décisions de systèmes à des références afin d'estimer la performance, la justesse et la robustesse des systèmes. Les systèmes sont évalués dans les mêmes conditions pour réaliser une tâche définie par les industriels pour laquelle les équipes de développeurs proposent des solutions. Les challenges répondent à un plan expérimental précis : un cas d'usage défini, un cadre expérimental réaliste (hardware in the loop, données, plateformes…), des métriques adaptées et des facteurs de variation à étudier identifiés.

Les références sont établies selon la nature des tâches, elles peuvent correspondre à des annotations humaines quand il s'agit d'interprétation, à des mesures physiques sur étalons ou à des actions attendues dans le cadre de la robotique. En fonction de la maturité des technologies, les cas d'usage traités (et leurs références) peuvent être fournis par les industriels ou en open data lorsqu'un organisme public a financé l'obtention des références. Les références peuvent également être accessibles par simulation, une fois cette dernière validée par des tests réels. Il est de la responsabilité de l'évaluateur de garantir la pertinence des références.

La métrique de comparaison est une fonction de coût que les développeurs cherchent à optimiser. C'est cette mesure qui permet de qualifier la capacité de l'IA à réaliser la tâche pour laquelle elle a été créée. Cette métrique est le critère qui permet de juger de pertinence du système d'IA.

Pour évaluer équitablement et simultanément la fiabilité des différentes IA, les challenges sont organisés par un tiers de confiance. La synchronie de la campagne est une solution pour garantir l'équité entre les systèmes d'IA.

Ces challenges peuvent facilement avoir une portée internationale s'ils sont correctement organisés et pertinents en termes d'application. Ils permettront, non seulement le benchmark des solutions, mais aussi d'assurer la visibilité et le rayonnement de la France.

---

[2] ROADEF : https://www.roadef.org/roadef-le-challenge
[3] TRAD : www.trad-campaigns.org, IWSLT initiative européenne
[4] MAURDOR : www.maurdor-campaigns.org, initiative française
[5] REPERE : www.defi-repere.fr, initiative française
[6] TAC: http://tac.sics.se/page.php?id=1 (Trading Agent Competition), PTAC: http://www.powertac.org/node/4 (Power Trading Agent Competition)
[7] ART: http://megatron.iiia.csic.es/art-testbed/ (Agent Reputation and Trust)
[8] Robocup: http://www.robocup.org/ , initiative internationale, http://rfia2016.iut-auvergne.com/index.php/ia/competition-ia-sur-robots, initiative française, http://www.enib.fr/~even/BotContest2017.html , initiative française
[9] https://multiagentcontest.org/, initiative internationale
[10] https://www.udacity.com/self-driving-car, initiative américaine

2.1.1.4 Développement de plateformes interopérables : faciliter la mise en œuvre des tests

Pour la mise en œuvre de ces challenges et permettre aux développeurs de situer leurs développements par rapport à des challenges précédents, des plateformes d'évaluation sont nécessaires. Elles permettent de garantir la reproductibilité des expérimentations pour que tous les développeurs puissent réaliser la même expérience depuis leurs locaux. Il s'agit de définir une interface claire et stable qui contient tous les résultats des challenges précédents. Ces plateformes sont développées et maintenues en open source. L'organisateur des challenges est garant de l'intégrité de la plateforme.
En fonction des applications, les plateformes peuvent inclure des éléments physiques de type *hardware in the loop*, des simulateurs ou bien des bases de données conséquentes.

Le développement de la plateforme doit s'appuyer sur les standards existants afin d'assurer l'interopérabilité des composants des systèmes proposés.

2.1.1.5 Développement de chaîne d'outils à disposition des industriels

Afin d'améliorer les performances de systèmes complexes, il est essentiel de développer des chaînes d'outils, probablement à partir de solutions en open source, pour la conception des algorithmes et des systèmes. Cela permet aux industriels qui en disposent de pouvoir s'appuyer, comme dans d'autres domaines (ingénierie dirigé par les modèles, compilation sur des architectures matérielles, etc.), pour assurer la qualité non seulement des systèmes, mais aussi des processus sur lesquels ils 's'appuient. Ces derniers sont indispensables pour aller vers la garantie et la certification.

2.1.2 Propositions de dispositifs et moyens : données, connaissances et plateformes matérielles et logicielles permettant aux entreprises de tester les technologies issues de travaux de recherche

Face aux moyens considérables, tant financiers qu'en terme d'infrastructures, investis par les « majors » » américains de l'IA, l'ensemble des acteurs français du domaine s'accordent sur la nécessité **d'efforts massifs de mutualisation**.
En effet, afin de pouvoir mettre à disposition de nos industriels des algorithmes performants et compétitifs, nos chercheurs ont besoin de s'appuyer sur des volumes de **données** dont aucun acteur ne dispose à lui seul. Il en va de même en ce qui concerne les **infrastructures** de calcul dédiées à l'IA, extrêmement consommatrices.
Mais les besoins de mutualisation et d'animation concertée sont encore plus larges. Ainsi, une même attention est à porter aux enjeux d'« exposition », de communication et de médiation sur la disponibilité de tels ou tels résultats opérationnels de recherche, assortis des cas d'usage auxquels ils savent répondre et ainsi que des **connaissances** qui en émergent.
Dans ce contexte, la question de l'échelle d'intervention souhaitable s'est posée au sein du groupe qui vise dans ce chapitre à présenter, pour l'ensemble des besoins, les pistes de solutions adaptées au niveau local, national et européen.
Comme le soulignent les dernières positions du Parlement européen, sur les questions de robotique et d'intelligence artificielle, l'ensemble des acteurs publics et privés européens a besoin de partager une expertise technique, éthique et réglementaire, et en particulier dans un environnement international qui appelle à traiter sérieusement les questions de souveraineté quant aux choix des outils d'intelligence artificielle sur lesquels investir.

2.1.2.1 Plateformes génériques

Les acteurs de l'IA français ont besoin de pouvoir s'appuyer sur **une vitrine et une marque nationales** valorisant l'excellence de la recherche via ses publications, logiciels et données dans ces domaines tout en facilitant leur accès pour les industriels, depuis la prise d'information sur les solutions existantes jusqu'à leur exploitation et leur mise en œuvre contractualisable.

La Commission européenne s'orientant actuellement vers une plateforme « **AI On-demand** », il est recommandé à tous les laboratoires de recherche d'assurer une présence massive de leurs travaux sur cette plateforme qui sera d'ores et déjà financée et mutualisée, avec un fort potentiel de fédération pour les partenaires français et européens. Des mesures incitatives pourraient être développées par les pouvoirs publics qui peuvent, tout au moins, assurer la promotion de ce futur dispositif auprès de leurs différents réseaux.

Cependant, afin de pouvoir tirer parti de ce type de plateforme européenne à venir et de s'assurer que les acteurs français y bénéficieront de toute la visibilité nécessaire, il est également recommandé d'y contribuer sous la forme d'une **plateforme nationale**, celle de l'Europe étant appelée à être une sorte de « plateforme de plateformes ».

Une telle **plateforme « #FranceIA »** serait adaptée à des applications industrielles, pour des solutions d'intelligence artificielle en général et non pas seulement en data science. Celle-ci aurait comme principaux avantages de :

- permettre d'accéder à un ensemble de données et algorithmes, travailler à leur mise en forme ou leur enrichissement ;
- expérimenter de nouveaux produits et services, usages, modèles d'affaires ou de nouvelles briques technologies, y compris pour la sécurisation des échanges et des données ;
- travailler à l'interopérabilité des données et à la régulation des échanges entre partenaires (cf. « Industrial Data Space » ou « free flow of datas ») ;
- fédérer les acteurs, qu'ils soient utilisateurs, fournisseurs de service ou de technologies, organisme de recherche, etc. ;
- impliquer les industriels dans la production de connaissances expertes à partir desquelles les chercheurs pourront travailler, telles que des ressources terminologiques / thesaurus / ontologies / vocabulaires.

A noter aussi qu'une telle plate-forme pourrait aussi être utile pour la mutualisation des algorithmes pour SAT (satisfaisabilité) et ASP (*answer set programming*) sur lesquels nous avons une excellente communauté de recherche.

2.1.2.2 Plateformes thématiques prioritaires

Trois plateformes prioritaires sont proposées :

- **Données, logiciels, ressources de calcul pour l'apprentissage automatique**
  Cette recommandation vise à assurer aux acteurs français de l'écosystème IA la mise à disposition des capacités de calcul nécessaires, indispensables pour être crédible dans la compétition internationale.
  La plateforme regroupera des ensembles de données (vision, parole, industrie, marketing, banque/assurance, etc.), une librairie d'algorithmes open source basée sur l'outil scikit-learn, et sera dotée de moyens de calcul surpuissants en particulier à base de GPU, architecture matérielle adaptée aux algorithmes d'apprentissage automatique.
  Les moyens de calcul pourront être portés par GENCI.

- **Plateforme véhicule autonome** (en relation avec le GT correspondant) ;
  Sur le plan matériel, cette plateforme comprendrait l'investissement dans des pistes d'essai, sites et véhicules ; sur le plan logiciel, les simulateurs et boîte à outils.
  Une quantité colossale de codes et de données en open source sponsorisés par les grands groupes (Udacity, Google, OpenAI) est disponible. Il s'agirait également d'y participer en accompagnant les initiatives des constructeurs français pour les positionner au mieux sur le secteur. En effet, le soutien du secteur nécessite un co-investissement public-privé conséquent.

- **Ressources pour le traitement automatique de la langue naturelle, l'interaction humain-machine et les agents conversationnels**
    - Le développement d'agents conversationnels est identifié comme priorité par les entreprises dans le cadre de la relation client. Les ressources de traitement de la langue naturelle en sont un élément clé, en particulier pour ce qui concerne la langue française sur laquelle notre pays a un rôle fédérateur à jouer.
    - Ouverte à tous secteurs d'activité, la plateforme de ressources mettra à disposition des acteurs des textes annotés, traduits, étiquetés, etc. Elle permettra également aux entreprises utilisatrices de réaliser des tests d'agents conversationnels.

2.1.2.3 Autres plateformes thématiques

- **Plateforme d'intégration et simulation**
  Plateforme matérielle et logicielle, par exemple autour des « architectures neuromorphiques ». Elle aurait pour objectif de développer toute la chaîne d'outils nécessaires, allant de la conception d'applications à l'implémentation adaptés aux matériels, tenant compte des contraintes de consommation, de puissance du matériel, etc.

- **Plateforme d'intégration multi-agents et Internet des Objets**
  D'un point de vue multi-agent (plus généralement de la coopération, coordination d'agents encapsulant des technologies développées dans divers domaines de l'Intelligence artificielle), il y a besoin d'une plateforme pour le parangonnage à grande échelle dans un cadre ouvert, hétérogène et décentralisé d'algorithmes de coordination et de régulation d'agents autonomes : agents développés par des opérateurs différents, déployés dans un environnement dynamique et distribué en interaction avec besoin de satisfaire des buts communs ou non.
  Une extension à l'Internet des Objets est pertinente dans la mesure où l'approche multi-agents permet de traiter les questions de communication, de coordination, de prise de décision collective pour des objets « lourds » dotés de capacités de raisonnement.
  Il sera possible de décliner cette plateforme sur différents domaines applicatifs (énergie, transport, véhicules autonomes, pilotage d'ateliers de production, etc), voire de les combiner (p.ex. pour la ville intelligente).

2.2 Approfondissements sur les thèmes de l'E-éducation, de la Robotique, de la Santé et des applications médicales et de l'Energie

2.2.1 Synthèse des approfondissements thématiques

E-éducation

Dans un monde dans lequel les questions d'éducation deviennent centrales et où la soif de connaissance et la demande de formation explosent, l'enseignement en ligne est une modalité pédagogique en plein essor. Les technologies d'intelligence artificielle permettent de mieux gérer les processus d'apprentissage, d'en adapter les contenus ; elles facilitent la personnalisation des parcours, notamment grâce à l'analyse des traces des apprenants. Notre pays compte plusieurs jeunes pousses dans ce domaine, des projets institutionnels avancés, des équipes de recherches de niveau international et une certaine avance sur le plan éthique. Au vu des enjeux économiques et sociétaux de l'enseignement en ligne, il est important de soutenir ce domaine et en particulier de poursuivre le transfert des résultats de recherche vers les entreprises et vers les plateformes de MOOC.

Robotique

Avec une forte culture scientifique et d'ingénierie, la France a une excellente recherche en robotique. Même si elle n'a pas encore de filière industrielle structurée dans le domaine, elle possède des PME très compétitives au niveau international. Les techniques d'intelligence artificielle sont indispensables au bon fonctionnement des robots : *deep learning* (pour la reconnaissance d'objets, de visages, d'activités, de sons, de parole, etc.), lien entre perception et action, apprentissage de notions de sens commun et de nouvelles compétences tout au long de la vie du robot, généralisation de concepts appris dans une situation et réutilisables dans une autre, etc. Dans ce contexte, les données constituent un goulot d'étranglement. Afin de ne pas devenir dépendant des acteurs multinationaux maîtrisant les grandes masses de données et de ne pas se laisser distancer dans le domaine de la robotique, il convient que les grands industriels français se fédèrent pour alimenter des corpus de données collectés et indexés (images, vidéos, sons, etc.) afin que les laboratoires de recherche puissent y puiser. Par ailleurs, les contraintes réglementaires françaises très restrictives quant à l'utilisation des données peuvent constituer un frein au développement des travaux dans le domaine.

Il est donc recommandé de revisiter ces politiques publiques dans une optique d'assouplissement des contraintes en question.

Enfin, une implication plus forte est souhaitable dans l'élaboration des normes, notamment relatives aux problématiques de sécurité.

Santé et applications médicales

De par l'excellence de ses équipes académiques et industrielles dans les disciplines qui concourent au domaine du numérique pour la santé et la médecine, la France a la capacité de créer de véritables « champions » internationaux. Par ailleurs, la quasi-totalité des champs de l'intelligence artificielle peut être convoquée par les applications dans ce domaine : de la représentation des connaissances et la modélisation du raisonnement à la robotique, en passant par l'apprentissage statistique et le traitement automatique du langage naturel. Or, force est de constater que les possibilités de transfert des travaux de recherche vers l'industrie sont considérablement freinées par rapport au fort potentiel et ce, pour

des raisons multiples : difficultés d'accès aux données, contraintes réglementaires, niveau faible des financements publics et privés par rapport à d'autres pays tels que les Etats-Unis par exemple, disponibilité des chercheurs, etc.
Plusieurs recommandations sont formulées dans la perspective de desserrer ces freins, parmi lesquelles on peut citer : le développement de l'approche française *Open data* en favorisant la constitution de bases de données médicales et de santé (et permettre aux entreprises françaises d'y accéder en priorité ?), la création de chaires mixtes dans les hôpitaux, la capacitation de la CNIL pour qu'elle puisse donner des avis plus rapides sur la mise en place de flux de données médicales accessibles aux chercheurs et aux entreprises, la mise en place d'un fonds d'investissement dédié aux technologies d'IA pour la santé, développement de la télémédecine, la formation aux nouveaux métiers impliquant la mise en œuvre du numérique et de l'IA dans le domaine de la santé et l'actualisation des formations en santé avec les nouveaux outils de l'IA.

Energie
Les énergies renouvelables sont aujourd'hui arrivées à « parité réseau » et le développement puis l'adoption par le monde industriel de l'intelligence artificielle font de l'IA un élément clé de la quatrième révolution industrielle. On pourra optimiser les processus de maintenance et de production et réaliser des économies d'énergie. On pourra également planifier et piloter la transition énergétique grâce à des décisions guidées par la donnée.
L'IA peut être le moteur des solutions aux grands enjeux :

1. Afin de révéler et activer l'efficacité et la flexibilité de la demande, sans impacter l'usage, l'IA peut permettre de dégager des profils de consommation efficace et flexible voire d'activer certaines consommations, sans "dégrader" l'usage. Elle peut également promouvoir la recherche de solutions décentralisées et coopératives d'un équilibre entre les besoins consommateurs et la stabilité du réseau électrique.
2. Afin de traiter le problème de l'équilibrage d'un réseau à maille fine, imposé par les sources renouvelables distribuées et intermittentes, il serait envisageable de concevoir un système avec de nombreux systèmes d'IA, maille par maille. L'IA ouvre des perspectives très prometteuses pour permettre, de manière récursive, l'équilibrage depuis les mailles les plus fines jusqu'au système national, néanmoins, la question se pose d'une IA souveraine.
3. Pour structurer un marché de l'électricité qui reflète la stratégie nationale et européenne, l'intelligence artificielle peut aider au design de structure de marchés, plus subtil, plus local, et proposer un nouveau support décentralisé à la décision.
4. Un accès contrôlé aux données peut contribuer à promouvoir la filière, c'est également un prérequis au déploiement de l'IA dans le cadre du réseau électrique. Le groupe recommande une approche raisonnée de visibilité adaptée à chaque cas d'usage.

En, synthèse, le groupe recommande le développement de deux prérequis – une **matrice de visibilité de l'ensemble des données du réseau** et le développement d'un **avatar temps-réel du réseau** – ainsi que la mise en place de deux Grands Défis : le développement d'une **solution d'équilibrage**, et la conception de **places de marché pour l'énergie**.

2.2.2 Détails des approfondissements thématiques

2.2.2.1 E-Education

L'enseignement en ligne : une opportunité industrielle pour l'intelligence artificielle

Après avoir assisté les humains pour effectuer des tâches manuelles, les machines les assistent désormais pour effectuer également des tâches intellectuelles : ce ne sont donc plus uniquement les métiers de tisserand, de potier, de fraiseur, etc. qui se transforment, mais également, ceux de médecin, de juriste, d'ingénieur, etc.

Le métier d'enseignant se transforme, lui aussi, en profondeur : l'enseignement en ligne, issu de la fusion de l'enseignement assisté par ordinateur, de l'enseignement à distance et de l'enseignement mutuel, s'impose, peu à peu, dans de nombreuses formations. Certaines plateformes comptent plusieurs millions d'inscrits et, même s'il est très difficile d'évaluer précisément l'impact mondial de l'enseignement en ligne, il semble que près de soixante millions d'apprenants se sont inscrits à une formation en ligne en 2016.

Nota : dans cette partie, les termes « apprentissage », « apprenants » désignent les activités et les apprenants humains utilisant l'enseignement en ligne. Les termes « *apprentissage automatique* » désignent les algorithmes technologies permettant à des machines de faire évoluer leur connaissance en fonction de données d'observation des systèmes d'enseignement en ligne. Pour clarifier cette distinction, *apprentissage automatique* est toujours en *italique* dans le texte.

Les enjeux de l'enseignement en ligne

Les enjeux de l'enseignement en ligne sont multiples.

L'accès de tous à la connaissance. L'enseignement en ligne permet de donner un accès à la connaissance à un très grand nombre de personnes, même dans les pays qui ne disposent pas d'un très grand nombre d'enseignants. Cet accès est rendu plus facile encore par le fait que beaucoup de formations sont désormais accessibles, non seulement sur des ordinateurs, mais aussi sur des téléphones.

La formation tout au long de la vie. L'enseignement en ligne permet aux apprenants de se former tout au long de leur vie, sans contrainte horaire et sans barrière sélective.

De nouvelles formes de remédiation. L'enseignement en ligne permet de nouvelles formes de remédiation pour les apprenants en difficulté, en adaptant le rythme de l'enseignement et le niveau des exercices à chacun. Il porte ainsi l'espoir de changer le destin des 110 000 jeunes qui quittent chaque

année le système scolaire sans diplôme et dont le coût, cumulé sur l'ensemble de leur vie professionnelle, est évalué à 230 000 € chacun[11], soit 25 milliards d'€ par an au total.

La compréhension des processus d'apprentissage. Les plateformes d'enseignement en ligne constituent un formidable outil pour observer la manière dont nous apprenons : l'efficacité d'un cours ou d'un exercice peut, par exemple, être observée sur des milliers d'apprenants. Elles constituent ainsi une solution partielle au problème de la difficulté de l'observation dans les sciences de l'éducation et dans les sciences humaines en général. Elles permettent également, de par la collecte et l'analyse des données d'apprentissage sur les apprenants, d'améliorer les méthodes d'enseignements grâce à des feedbacks automatiques sur l'efficacité des contenus pédagogiques. Cette collecte pose de nouvelles questions éthiques : observer, comprendre et agir dans le respect des apprenants nécessite une instrumentation innovante et inclusive, constituant un défi que les entreprises commencent à relever.

L'organisation des systèmes éducatifs. L'enseignement en ligne nous mène à repenser notre système éducatif, que la société industrielle a façonné en lui donnant une organisation pyramidale – le professeur a un supérieur hiérarchique : l'inspecteur, qui a lui-même un supérieur hiérarchique : le recteur, etc. –, qui répond à l'organisation pyramidale des entreprises. Le découpage entre enseignement général, technique et professionnel répond également à la division du travail dans ces entreprises. La transmission de l'information d'un vers plusieurs répond à la transmission de l'information dans ces entreprises. L'enseignement en ligne permet, au contraire, une organisation en réseau, il décloisonne les enseignements et il permet à chacun d'être, tour à tour, enseignant et apprenant.

Une opportunité industrielle. L'enseignement en ligne est un enjeu économique. Notre pays compte de nombreuses jeunes pousses (*start-ups*) dans ce domaine : OpenClassroom, Éduclever, Gaya Technologies, NeuralCat, Unow, Hoomano, 360 learning, etc. qu'il importe d'aider à devenir des leaders mondiaux, qui favoriseront, à leur tour, l'émergence de nouvelles pousses.

Un outil au service de la performance des entreprises. En créant des formations en interne (*in-house*) en ligne accessibles à tous les salariés, les employeurs ont la possibilité d'associer l'apprentissage individuel aux besoins organisationnels de l'entreprise. Ils peuvent ainsi concevoir des systèmes de formation utilisant des équipements et des technologies qui font déjà partie des processus organisationnels, et des modules d'information et d'apprentissage répondant aux besoins de leurs employés.

La souveraineté. L'enseignement en ligne est un enjeu de souveraineté nationale. Les leaders mondiaux de l'enseignement en ligne promeuvent, volontairement ou non, les systèmes éducatifs et les méthodes pédagogiques de leurs pays et de leurs cultures. Nous ne devons pas nous laisser imposer ces systèmes éducatifs, sans promouvoir également le nôtre. Par ailleurs, les apprenants qui utilisent une plateforme d'enseignement en ligne y laissent des traces : profils, parcours d'apprentissage, évaluations, questions, etc. dont nous ne pouvons pas laisser l'exploitation à des entreprises étrangères dans des conditions éthiques que nous ne maîtriserions pas.

---

11 http://www.gouvernement.fr/action/le-decrochage-scolaire

Les MOOCs, à la recherche de modèles économiques viables. La première raison d'être des *Massive Open Online Courses* (MOOC – Cours en ligne ouverts et massifs) est de démocratiser l'enseignement supérieur en ouvrant l'accès à la connaissance à tout un chacun gratuitement. Ces cours se sont développés de manière exponentielle ces dernières années et rassemblent des millions d'apprenants. De ce fait, même s'ils ont un coût de développement initial assez élevé, ils permettent aux institutions de les proposer gratuitement ou à faible prix. Cette gratuité des formations rend cependant difficile la production de valeur ajoutée économique à partir du seul contenu. Les plateformes d'hébergement des MOOC développent ainsi d'autres formes de monétisation qui, pour une grande part, ne reposent pas sur les apprenants : de la vente de certificats (dont le prix varie selon qu'ils concernent des cours à visée professionnelle ou pas), de la proposition de services *employee-matching* (mise en relation entre les entreprises et les apprenants dont le coût est entièrement supporté par les entreprises), vente des droits d'utilisation des MOOC à d'autres institutions. Certaines plateformes, telle que *France Université Numérique* (FUN), misent également sur le marché de la formation continue. Elles proposent des SPOC – *Small Private Online Courses* – qui sont des cours en ligne sur le même format que les MOOC mais conçus pour un nombre plus restreint d'apprenants afin d'assurer un encadrement plus efficace. Elles sont généralement adaptées aux besoins des entreprises.

Les apports de l'intelligence artificielle

Les premières plateformes d'enseignement en ligne se contentaient de délivrer des cours, sous forme de vidéos, et de proposer des exercices aux apprenants, sous forme de questionnaires à choix multiples. Les méthodes issues de l'intelligence artificielle sont essentielles pour rendre ces enseignements plus diversifiés, plus adaptatifs et plus efficaces.

La diversité des exercices. L'apprentissage passe par la réalisation d'exercices, telle la réponse à des questions fermées ou ouvertes. Par exemple, après avoir étudié la révolution de la Terre et les saisons, un apprenant peut se voir proposer une question fermée « Où fait-il plus chaud en décembre ? », suivie de plusieurs réponses possibles : en Afrique du Sud, au Canada, ou en France. Mais les mêmes connaissances peuvent être testées de manière plus approfondie par question ouverte « Pourquoi, en décembre, fait-il plus chaud en Afrique du Sud qu'au Canada ? ». Cette dernière question est pédagogiquement plus intéressante, car, dans sa réponse, l'apprenant doit mentionner l'inclinaison de la Terre par rapport au plan de l'écliptique, là où, dans la question à choix multiples, il lui suffisait de cocher une case. Aujourd'hui, ces questionnaires sont la plupart du temps réalisés par les enseignants eux-mêmes. Pour améliorer leur efficacité, les enseignants doivent être assistés dans leur construction. Il est en particulier essentiel d'engendrer des exercices de difficultés différentes, afin de proposer des exercices adaptés au niveau de chaque apprenant.

La correction automatique des questionnaires à choix multiples est beaucoup plus simple que celle des questions ouvertes, qui nécessitent d'analyser la réponse rédigée par l'apprenant. Une démarche est l'organisation d'une évaluation entre pairs, qui, pour évaluer les réponses d'un apprenant, exploite les capacités intellectuelles d'un autre apprenant. Une autre démarche est l'utilisation de correcteurs automatiques de réponses ouvertes exploitent des techniques *d'apprentissage automatique*, dans lesquelles un algorithme d'intelligence artificielle doit « comprendre » la réponse fournie par l'apprenant pour en évaluer la pertinence. De tels algorithmes existent, mais ils demandent à être améliorés et déployés à grande échelle.

Une meilleure indexation des ressources. L'efficacité d'un processus d'enseignement en ligne dépend beaucoup de la capacité à proposer à chaque apprenant une liste de cours qui correspondent à son

niveau, à ses aspirations et à ses projets. Pour cela, il est essentiel que les ressources pédagogiques soient référencées et décrites précisément. Des outils de recommandation, fondés sur des techniques *d'apprentissage automatique*, sont déjà utilisés sur les plateformes de commerce électronique ou de vidéo à la demande. Ils demandent à être développés et adaptés aux plateformes d'enseignement en ligne pour permettre aux apprenants de mieux s'orienter dans cette nouvelle offre de formation.

Une meilleure efficacité dans la gestion des questions. Quand une formation en ligne s'adresse simultanément à plusieurs dizaines de milliers d'apprenants, le nombre de questions posées par ces apprenants peut se compter en milliers. Il faut alors trouver des moyens pour y répondre. Les systèmes *d'apprentissage automatique* peuvent aider les enseignants à répondre, voire répondre à leur place, en exploitant des réponses passées à des questions similaires. Ainsi, plus les enseignants répondent, plus le système apprend à répondre à leur place et plus il les aide. Celui-ci peut alors se concentrer sur les questions les plus complexes sur lesquelles son expertise est utile. Il est aussi possible de mettre en place un système d'inter-tutorat, pour un enseignement mutuel entre les apprenants les plus avancés et les moins avancés sur un même cursus. Quand la réponse aux questions est fournie, non automatiquement, mais par les autres apprenants, les systèmes *d'apprentissage automatique* permettent de regrouper les questions par thème, de classer les réponses, des détecter les situations nécessitant des interventions complémentaires, etc.

L'analyse des traces laissées par les apprenants. En construisant des profils, en choisissant des parcours, en répondant à des évaluations, en posant des questions, etc., les apprenants laissent des traces sur les plateformes. Ces traces constituent des observations essentielles pour comprendre la manière dont les humains apprennent, mais aussi la manière dont chaque apprenant singulier apprend. Modéliser chaque apprenant en analysant ses traces, permet par exemple de détecter les apprenants en risque de décrochage. Cela permet aussi de proposer aux apprenants une formation sur mesure (en fonction de leurs préférences, de leurs compétences, de leurs rythmes et méthodes d'apprentissage, etc.), même quand l'offre de formation est massive. Cela permet enfin de former des groupes de travail judicieux, en construisant des groupes de niveaux homogènes ou au contraire en mélangeant les niveaux pour favoriser le parrainage. Tout cela demande cependant de savoir analyser ces traces qui sont souvent volumineuses. Il faut, pour cela, encore utiliser des techniques *d'apprentissage automatique*.

Au-delà des indicateurs automatiquement calculés sans le concours des acteurs de l'apprentissage, les techniques d'intelligence artificielle permettent de considérer l'observation des processus d'apprentissage comme une source de co-construction de connaissances entre apprenants, tuteurs, enseignants, formateurs, etc. et ceci tout au long de la vie. L'activité d'apprentissage peut en effet être l'objet d'une documentarisation sémantique dès l'instant où les traces d'apprentissage sont modélisées, au sens d'un modèle de connaissance, progressivement par transformations explicites permettant d'intégrer les points de vue de tous les acteurs et de créer des espaces de négociation pour dégager des points de vue commun. Ces traces modélisées d'apprentissage peuvent alors être gérées par les acteurs de l'apprentissage au premier rang desquels les apprenants, puis mises à disposition des autres acteurs, en particulier les chercheurs pour l'élaboration d'indicateurs nouveaux, afin de donner du sens partagé aux processus d'apprentissage. Les entreprises qui fournissent les environnements de médiation des activités d'apprentissage peuvent offrir alors un service particulièrement innovant, permettant de documentariser, capitaliser, partager, réutiliser les pratiques d'apprentissage dans toute leur dynamique et non réduite à des indicateurs statiques dont l'usage peut s'avérer très réducteur hors contexte.

Une solution de taille pour la mise à jour des contenus pédagogiques. L'atout de l'intelligence artificielle est lié à sa capacité à évaluer, apprendre et adopter une stratégie dynamique. Le développement des enseignements en ligne est coûteux en temps : de 50 heures à plus de 200 heures pour la conception d'une heure de contenus en ligne contre un maximum de 82 heures pour celle d'une heure de cours en présentiel[12]. L'usage de l'intelligence artificielle peut atténuer ce coût de développement grâce à une classification et une organisation des données qui permettraient d'identifier des contenus transversaux à plusieurs thématiques. Par ailleurs, cet usage permet de mettre à jour les matériels pédagogiques de manière continue et d'identifier de manière automatique les modifications à apporter aux contenus.

Une amélioration des taux de rétention. Il s'agit là d'un des objectifs clé de la mise en place des *learning analytics*, surtout lorsque les frais d'inscription constituent une part importante du financement de certaines plateformes d'enseignement en ligne. Il est en effet beaucoup plus facile de retenir un apprenant déjà inscrit que d'en recruter un autre sur un marché de plus en plus compétitif. Les données sur la manière dont les apprenants interagissent avec les contenus pédagogiques peuvent être un indicateur de leur engagement et, par conséquent, de la probabilité qu'ils décrochent. Les systèmes analytiques permettent d'identifier les apprenants qui se désengagent très tôt dans le parcours et de choisir le mode d'intervention le plus approprié afin d'éviter leur décrochage.

Exemples de projets industriels

Gaya Technologies. *Intelliquiz* est un moteur de génération automatique de questionnaires à choix multiples à partir de bases de données et du Web, développé par la société Gaya Technologies. Il permet de créer des questionnaires de façon infinie, dans toutes les langues et pour tous les programmes scolaires. Il s'adapte au profil de l'apprenant. Il utilise des techniques issues des jeux vidéo, du Web sémantique et de l'intelligence artificielle. Il est fondé sur la génération automatique de réponses alternatives, la génération automatique de questions, l'indexation des questionnaires dans un référentiel de connaissances et de compétences et l'adaptation au profil de l'apprenant.

Le développement d'un tel projet industriel mène à poser de nombreuses questions de recherche : Comment extraire à partir d'un jeu initial de questions en langue naturelle, une représentation de ces questions dans les formalismes du Web sémantique ? Comment extraire du Web de données des réponses fausses mais crédibles à des questions formalisées ? Comment évaluer la qualité des documents utilisés ? Comment générer des questions à partir d'un ensemble de critères données – notamment des thématiques –, en langue naturelle et formalisées dans les modèles du Web sémantique ? Comment construire une séquence de questions à proposer à l'apprenant ? Quelles méta-connaissances sur les questions sont nécessaires pour cela ? Comment adapter une séquence de questions à l'utilisateur ? Quelles connaissances sur l'utilisateur sont nécessaires pour cela ?

---

[12] Chiffres issus d'une enquête de Bryan Chapman (Chapman Alliance) en 2010 auprès de 249 organismes dans le monde ayant créé du contenu pédagogique suivi par près de 20 millions d'apprenants (enseignement en ligne, apprentissage mixte, présentiel)

Chapman, B. (2010), *How Long Does it Take to Create Learning?* http://www.chapmanalliance.com/howlong/

De tels problèmes doivent être résolus par une coopération entre ingénieurs et chercheurs.

Éduclever. L'entreprise Éduclever développe des services multiplateformes qui reposent sur trois fondamentaux : une ontologie (modélisation par une représentation informatique) des compétences et connaissances liées aux programmes scolaires, un référentiel praxéologique (liste d'objectifs d'apprentissage organisés en savoirs et savoir-faire) et la construction en temps réel d'un profil de chaque apprenant, qui concerne ses connaissances, ses compétences, ses conceptions, son comportement, etc. L'association de ces trois axes permet de proposer des activités personnalisées qui donnent à chacun la possibilité d'avancer à son propre rythme et d'éviter de décrocher.

Le développement d'un tel projet industriel mène à poser de nombreuses questions de recherche : Comment construire un référentiel précis et complet des connaissances ? Comment indexer les ressources pédagogiques ? Comment aider les utilisateurs en ligne ? Comment aider les auteurs à créer des stratégies pédagogiques ? Comment générer automatiquement des exercices ?

Ces questions font l'objet de coopérations entre l'entreprise et des équipes de recherche (TWEAK - LIRIS, METAH - LIG, WIMMICS-Inria), matérialisées notamment par des thèses CIFRE.

NeuralCat. NeuralCat a été créée pour aider aux partage et diffusion de la connaissance. Poursuivant ce but, elle développe et distribue une intelligence artificielle conversationnelle. Cette intelligence artificielle s'incarne sous la forme d'un chat virtuel qu'il faut entraîner. Le chat est capable de syllogisme et demande le cas échéant à l'utilisateur-entraîneur de confirmer ses déductions, ce qui oblige celui-ci à s'assurer de la validité de son enseignement. Une fois ses classes faites, le chat se confronte aux chats des autres utilisateurs-entraîneurs à travers des compétitions de connaissance ou d'intelligence. Ce renversement de la flèche de l'apprentissage permet d'éviter le premier des écueils, connu de tout enseignant : la perte de motivation.

NeuralCat vise les marchés du jeu vidéo familial et des écoles hors-contrat.

Exemples de projets institutionnels

France Université Numérique. Le projet France Université Numérique (FUN) est né en 2013, sous l'impulsion du ministère de l'Enseignement Supérieur et de la Recherche, qui a souhaité mettre à disposition des établissements français une plateforme de MOOC, afin de leur permettre de diffuser largement leur MOOC et gagner en visibilité à l'international, en particulier dans le monde de la francophonie. En trois ans, plus de 90 établissements ont produits 270 MOOC, répartis en plus de 460 sessions qui ont suscité plus de 2,6 millions d'inscriptions.

Depuis 2015, la plateforme est opérée un groupement d'intérêt public et de nouvelles activités se développent. L'usage des MOOC se diversifie. Ils contribuent à la rénovation des pratiques pédagogiques avec leur réutilisation sur les campus via l'émergence de la classe inversée et des dispositifs hybrides. Ils suscitent également un intérêt grandissant au sein des entreprises et des acteurs du monde socio-économique, pour la formation professionnelle. Le GIP accompagne ses nouveaux usages avec le déploiement de plateformes dédiées à ces nouveaux développements.

France Université Numérique coordonne le projet IDEFI-Numérique EIFFEL a qui bénéficie d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme d'Investissements d'Avenir, et travaille dans ce contexte avec plusieurs jeunes pousses innovantes, un laboratoire de recherche et dix établissements d'enseignement supérieur. Ce projet vise à rendre les MOOC plus

interactifs et accessibles en intégrant sur la plateforme des nouvelles fonctionnalités développées par les jeunes pousses partenaires, comme l'enrichissement de vidéos, les interactions en temps réel via des outils collaboratifs, ou encore l'ancrage mémoriel, première étape vers le déploiement d'outils d'intelligence artificielle permettant l'apprentissage adaptatif personnalisé.

Inria Learning Lab a pour ambition de faire de l'éducation numérique un enjeu de recherche majeur et de fédérer l'ensemble des équipes de recherche Inria qui abordent ces thématiques. Inria Learning Lab, et le Mooc Lab d'Inria qui l'a précédé, ont permis la production de quatorze MOOCs entre 2013 et 2016 - dix de ces MOOC ont été réalisés dans le cadre du sous-projet uTOP « Valorisation de la Recherche par la Formation ». Ces Mooc ont donné lieu à plus de vingt-deux sessions pour un total dépassant les 90 000 inscrits.

Inria Learning Lab apporte aux chercheurs Inria qui le souhaitent son expertise en matière de production et de réalisation de ressources pédagogiques. Il anime le réseau de chercheurs Inria en e-éducation à travers des journées scientifiques et sert de catalyseur entre les équipes pour monter des projets de recherche. Inria Learning Lab soutient également le développement de nouvelles technologies et de projets innovants pour l'e-éducation.

Exemples de projets de recherche

L'analyse de traces. Plusieurs équipes, notamment françaises – l'équipe TWEAK du LIRIS, l'équipe SICAL et plusieurs équipes à Grenoble, Le Mans, Brest, etc. – s'intéressent aux traces d'interaction comme un objet informatique formalisé, associé à des méthodes opératoires – transformations de traces – permettant de considérer toute trace ainsi modélisée comme le support d'interprétations formalisées et donc explicites. Le projet Hubble de l'Agence Nationale pour la Recherche est consacré à la question de l'observation de l'apprentissage humain à des fins d'analyse et d'exploitation (*learning analytics*).

Ces travaux sont maintenant disponibles pour un transfert efficace vers les entreprises. Un environnement qui articule des méthodes du Web et de l'intelligence artificielle, l'environnement *Kernel for Trace Based Systems* est partagé par plusieurs équipes de différents pays. Des partenariats avec des entreprises ont déjà été noués notamment pour l'assistance dans le cadre de l'apprentissage humain et pour l'assistance à l'analyse d'activité mais ces technologies – Web et intelligence artificielle – doivent être encapsulées de manière plus efficace dans les librairies des grands outils de développement du Web.

L'optimisation des parcours pédagogiques. La plateforme POEM (*Personalised Open Education for the Masses, http://poem.unistra.fr*) a comme objectif de créer des parcours pédagogiques au sein des multitudes de cours en ligne proposés. Cela permet d'offrir au plus grand nombre la possibilité d'apprendre en ligne, en formation initiale ou tout au long de sa vie, en construisant un parcours adapté. Cela permet également aux institutions de valoriser leurs contenus pédagogiques de manière automatique, en leur donnant une vie même après que les cours, exercices, vidéos proposés ne sont plus utilisés par ceux qui les ont conçus.

La première étape consiste à construire une carte de la connaissance, c'est-à-dire à représenter les contenus pédagogiques d'une manière ergonomique pour les apprenants, afin qu'ils puissent construire leur parcours, tout en répondant aussi aux problèmes d'orientation des apprenants. Les contenus pédagogiques sont organisés en sphères concentriques représentant des données reliées par des liens

verticaux (couches concentriques) et horizontaux (à la surface d'une couche). Pour choisir le niveau qu'il cherche à atteindre, l'apprenant peut zoomer vers l'intérieur ou vers l'extérieur de la sphère pour sélectionner la couche qui l'intéresse. Une fois le niveau choisi, il tourne la sphère pour trouver l'objectif qu'il souhaite atteindre. Il progresse ensuite dans la sphère des connaissances suivant un axe qui lui permet de traverser les disciplines de son choix.

La principale originalité de la plateforme POEM consiste à utiliser l'intelligence artificielle pour créer des trajectoires pédagogiques optimales en utilisant des algorithmes inspirés de la nature. En adaptant à l'être humain le fonctionnement des colonies de fourmis, le paradigme des *hommilières* a été validé avec succès par la société Paraschool sur plus de 500 000 élèves utilisateurs. Ces hommilières permettent aux apprenants de construire des chemins pédagogiques optimaux permettant d'atteindre des objectifs fixés le plus rapidement possible.

Dans le cadre de l'éducation, la force qui dirige les étudiants est l'évaluation qui permet aux apprenants de valider leur progression vers les couches supérieures de la sphère des connaissances jusqu'à l'obtention des certificats souhaités. L'évaluation dans la plateforme POEM ne s'effectue pas par des questionnaires à choix multiples mais par des questions ouvertes corrigées par d'autres apprenants, pour une évaluation formative de grande qualité pédagogique.

De même, un inter-tutorat entre étudiants plus ou moins avancés sur le même parcours est mis en place pour un apprentissage mutuel : les apprenants entrant dans le système ont des interlocuteurs humains à qui poser des questions, et en y répondant, les tuteurs (aussi des apprenants) acquièrent du recul sur les connaissances qu'ils ont acquises. Si un tuteur (lui-même apprenant) ne sait pas répondre à la question d'un autre apprenant moins avancé que lui sur le même parcours, il peut transférer la question à son propre tuteur. Si aucun des tuteurs ne sait répondre à la question, celle-ci parvient jusqu'à l'enseignant qui a proposé le cours ayant généré la question.

Questions éthiques

Les plateformes d'enseignement en ligne collectent donc de nombreuses données sur les apprenants : leurs profils, leurs parcours d'apprentissage, leurs évaluations, les questions qu'ils ont posées. Les développeurs et les enseignants doivent donc prendre en compte, dès la conception de ces plateformes et de ces cours, le fait que ces données sont des données privées que l'apprenant peut décider de diffuser ou non. Ce souci éthique est une exigence morale, mais c'est aussi un avantage concurrentiel, car il est peu probable que les apprenants acceptent d'utiliser des plateformes qui ne respecteraient pas le caractère privé de leurs données.

Les techniques de l'intelligence artificielle permettent de fournir les éléments théoriques et technologiques nécessaires à l'émergence de processus éthiques et des outils associés. Présentées plus haut, la gestion sémantique des traces d'apprentissage fournit le cadre théorique et technologique permettant *d'instrumenter* l'apprentissage en fournissant une panoplie de services innovants autour de la construction, la transformation, le partage, l'utilisation pour les indicateurs, la réutilisation comme bonnes pratiques, etc. Une démarche éthique par construction est alors possible et fournit une formidable plus-value potentielle à la médiation numérique des processus d'apprentissage.

Une opportunité industrielle

Dans un monde dans lequel les questions d'éducation deviennent centrales et où la soif de connaissance et la demande de formation explosent, l'enseignement en ligne est une modalité pédagogique en plein essor. Notre pays compte de nombreuses jeunes pousses dans ce domaine, des projets institutionnels avancés, des équipes de recherches de niveau international et une certaine avance sur le plan éthique. Il dispose donc, si ces projets sont soutenus, de véritables atouts pour prendre une place centrale, dans les pays francophones et dans le reste du monde, dans ce secteur industriel très dynamique sur le plan mondial.

2.2.2.2 Robotique

Enjeu/motivation/défis

La relation entre la robotique et l'intelligence artificielle est ambivalente : dans le contexte de la mission France IA, la robotique est vue comme une sous-partie de l'intelligence artificielle alors que les roboticiens ont plutôt tendance à voir une complémentarité avec l'intelligence artificielle, celle-ci comme étant une discipline bien plus jeune. Le robot s'inscrit en effet dans la longue histoire des machines : c'est une machine qui bouge et dont les mouvements sont contrôlés par un ordinateur. Il se distingue tout à la fois de l'automate dont les mouvements sont mécaniquement déterminés, et de l'ordinateur qui manipule des informations, mais ne bouge pas. Sans robotique, l'intelligence artificielle n'aurait pas d'action dans le monde physique, ce qui exclurait un énorme champ d'action (production industrielle, médecine, transports, assistance physique aux personnes, etc.).

Même dans la perspective « IA », la robotique doit donc être pensée comme ayant des ressorts propres. Le robot bouge pour agir. Il a affaire avec la physique du monde réel : le robot manipulateur doit pouvoir saisir un objet fragile sans le casser ou un récipient sans renverser son contenu ; le robot bipède doit garder son équilibre en se déplaçant. Manipulation et déplacement sont les deux composantes de l'action. C'est d'abord la maîtrise du mouvement des robots manipulateurs qui a permis le développement de la robotique industrielle. C'est ensuite la maîtrise du déplacement des robots mobiles à roues qui a permis dès les années soixante-dix d'envisager les grands programmes d'exploration planétaire. C'est aujourd'hui la maîtrise conjointe de ces deux fonctions motrices qui permet l'émergence de nouvelles machines capables tout à la fois de se déplacer et de manipuler. Et de même que le cerveau ne contrôle pas tout le corps, puisqu'il existe de nombreuses opérations inconscientes (e.g. digestion, régénération cellulaire…) voire mixte (respiration), ainsi les robots ne seront pas entièrement contrôlés par une intelligence artificielle.

Ceci met en lumière un défi important de la robotique face à l'IA : comment concevoir la relation entre l'IA et les robots ? Peut-on concevoir des « corps » qui puissent être finement contrôlé par une IA ? Peut-on ainsi délimiter les fonctions mécaniques et électroniques ? Y-a-t-il une pure « mécatronique robotique » qui se limiterait à une partie capteurs-actionneurs ? Est-ce que c'est faisable ? Est-ce que c'est efficace ? Une autre manière de poser le problème est de s'interroger sur la cohérence et sur la compatibilité entre les logiciels d'IA (et les modèles de données) et les parties robotiques (mécatronique) ; et plus généralement de s'interroger sur la caractérisation et sur l'analyse du couplage Robot-logiciel IA. Se pose alors aussi la question de l'adaptation des comportements générés par l''IA selon les évolutions des caractéristiques physiques du ou des robots (par exemple pour la gestion de la sécurité).

La résolution de ces enjeux est capitale, car c'est ce qui permettra d'avoir des robots plus agiles, plus adaptatifs, plus facilement coopératifs, et donc finalement plus productifs.

Plus spécifiquement vis-à-vis de l'IA, les techniques actuelles de *deep learning* sont nécessaires au bon fonctionnement des robots : reconnaissance d'objets, de visage, d'activité, de sons, de parole, etc. Les robots de service qui n'auront pas accès à ce type de fonction seront cantonnés à des rôles d'aspirateur. Les sociétés de robotique, comme SoftBank Robotics, ont les capacités intellectuelles pour intégrer ce type de technologies mais manquent de ressources pour les mettre en œuvre. La constitution des données pour nourrir le *deep learning* est un goulot d'étranglement et il faut se reposer sur les services (américains) des géants du domaine. Cela crée une dépendance risquée. En outre, au-delà du deep learning, les techniques d'IA doivent se développer pour intégrer, notamment pour les besoins de la robotique, des liens entre la perception et l'action, l'apprentissage tout au long de la vie du robot de nouvelles compétences, la généralisation de concepts appris dans une situation et réutilisable dans une autre, l'apprentissage de ce qu'est le sens commun, etc. Le *deep learning* est nécessaire au développement de l'industrie robotique mais la robotique aura besoin d'aller au-delà des capacités offertes par le *deep learning*.

SWOT

Forces

La recherche française est excellente en robotique.

La France a une culture scientifique et d'ingénierie très forte, qui lui donne des atouts en robotique.

Elle dispose de bons laboratoires (publics et privés) en IA et d'une longue tradition mathématique utile au développement de l'IA, et donc d'une bonne connaissance des besoins du domaine. Des grandes entreprises de télécommunication et de diffusion de contenu existent en France.

La France possède une culture diversifiée en IA – au-delà du Deep Learning – ce qui est un atout dans la filière robotique, puisqu'il n'est pas facile de déterminer a priori la discipline IA qui sera nécessaire à toutes les parties robotiques (p. ex. le contrôle/commande).

La France a d'excellente PME dans le domaine même si elle n'a pas de filière structurée.

Faiblesses

La France est historiquement faible en ce qui concerne la pénétration des robots industriels, probablement en raison d'une très mauvaise vision de l'impact de la robotique sur l'emploi. Pourtant, l'Allemagne maintient un taux emploi industriel plus élevé qu'en France, avec un taux de robotisation cinq fois supérieur.

L'apprentissage, en robotique comme en IA, requiert de très grandes bases de données. Or dès que ces données contiennent des images de personnes, la CNIL a une approche très restrictive. Ceci empêche la recherche en robotique tournée vers l'humain (cobotique, transport,…) de trouver une place en France. Il devient plus simple de faire cette recherche à l'étranger.

En outre, le stockage et le traitement de ces données pose problème, comme dit en introduction.

Avons-nous assez de centres de formation pour les experts en IA et pour des *data scientists* ? Avons-nous accès aux corpus de données qui permettent aux techniques d'apprentissage de déployer toutes leurs capacités ? Une certaine dispersion des forces est à noter.

Opportunités

Une prise de conscience du besoin d'agir émerge. La commission européenne va soutenir les développements dans ce domaine.

Il devient de plus en plus nécessaire d'intégrer des données logistiques (évolution de stock, préparation de batch, de commandes...) pour optimiser la gestion de la flotte globalement ou

localement (gestion de trafic) dans les entrepôts et au-delà. Cette question de compétitivité industrielle va forcer le couplage IA/Robotique.

La robotique (au-delà de l'automate industriel) prend son envol, ce qui va entraîner tout le secteur.

Menaces

Force est de constater un certain « siphonage » des compétences par les grands acteurs (GAFA) : Google développe un centre de 3000 chercheurs en IA à Zurich. En dépit de ses compétences, la France ne semble pas suffisamment attractive. Les jeunes formés sont plus attirés par la Californie ou la Suisse.

Aux Etats-Unis les entreprises recrutent à grande échelle (et à de bons salaires). Il y a accumulation d'un retard initial qui pourrait devenir irrattrapable.

Recommandations

La richesse est dans les données. Il faudrait que les grands industriels français alimentent des corpus de données collectées et indexées (image, vidéos, sons, etc.) pour que leur propre R&D ainsi que les laboratoires de recherche puissent y puiser. Le retard par rapport à des grands acteurs du domaine tels que Google est tel que si l'industrie française veut espérer ne pas se faire distancer complètement, elle doit se fédérer pour se donner les moyens d'obtenir des résultats probants. Cela passe peut-être par une mise en commun des moyens pour collecter et annoter ces données (pas seulement pour les stocker et les mettre à disposition).

Il est important aussi de tenir compte des normes (p. ex. ISO/TS 15066) et d'y contribuer (en influençant les orientations), en particulier pour les questions de sécurité.

2.2.2.3 Santé et applications médicales

Thèmes, enjeux et défis

Les équipes académiques et industrielles françaises ont des compétences importantes et reconnues en mathématiques, en informatique et en intelligence artificielle, mais aussi en ingénierie, en médecine et en santé. Associées à un système de soins bien structuré (CHU, sécurité sociale) et de grande qualité, dont les avantages doivent donc être mobilisés, à une bonne dynamique entrepreneuriale et à un bon système d'enseignement supérieur, elles ont la capacité de créer de véritables "champions" internationaux.

Elles doivent cependant faire face à des faiblesses, d'une part dans la gestion du temps (cette dimension devient essentielle pour réduire la durée des phases de maturation et de transfert), d'autre part lors de l'accès aux données, et enfin dans le financement tant amont (au niveau des chercheurs et des médecins) qu'aval (dans le soutien des entreprises). Les applications de l'IA dans le domaine de la médecine et de la santé s'appuient sur la plupart des domaines couverts par l'IA (Cf. rapport du GT recherche amont) :

- Représentation des connaissances et modélisation des raisonnements
- Apprentissage statistique et modélisation prédictive
- Décision et gestion de l'incertitude

- Satisfaction de contraintes et SAT (satisfaisabilité d'une formule logique)
- Planification et recherche heuristique
- Agents autonomes et systèmes multi-agents
- IA & web
- Traitement automatique des langues (TAL)
- Robotique
- Vision et reconnaissance des formes
- Modélisation cognitive et systèmes neuro-informationnels.

l n'est pas surprenant que nous retrouvions ici les mêmes défis, les mêmes enjeux et les mêmes thèmes que ceux qui ont été identifiés de manière plus générale par le GT recherche amont.

Il apparaît particulièrement important d'offrir un soutien large à l'ensemble des thèmes proposés, de favoriser des recherches multidisciplinaires, en portant attention aux interfaces (particulièrement importantes dans ces applications), ainsi que des recherches intégratives associant les industriels en prenant garde que leurs travaux nourrissent prioritairement le tissu socio-économique français.

On s'attachera à créer une dynamique autour des données : leur annotation, leur accès sécurisé, les plateformes de traitement, les codes et les ressources de calcul les enrichissant.

e document mentionne les spécificités de ces défis pour les applications médicales et de santé.

Ces points ont été mentionnés dans de nombreux rapports en France et à l'étranger sur la stratégie de la recherche en IA et dans le domaine de la santé (rapport stratégique de l'ITMO « Technologies pour la santé » d'AVIESAN, rapport du *National Science and Technology Council* américain, livre blanc d'Inria, cahier de veille de la Fondation Télécom et de l'IMT, rapports des autres groupes de travail de #FranceIA, etc.).

A la diversité des domaines cités s'appliquant aux données et connaissances de santé, s'ajoutent les dimensions éthiques et réglementaires dont il n'est pas possible de faire abstraction. Ces aspects ont en particulier un fort impact sur les possibilités de transfert. Le groupe a fait le constat que, si les laboratoires français sont très présents et réputés dans ces domaines, le transfert vers les industriels est vraisemblablement freiné par rapport au grand potentiel.

De même, le large potentiel de l'IA n'est pas encore pleinement exploité dès qu'il s'agit de faire des études à grande échelle, en partie à cause de la difficulté de l'accès aux données.

Apports et bénéfices attendus de l'IA dans le domaine de la médecine et de la santé

Ceux-ci sont nombreux et constituent des thèmes de recherche actuels et en pleine croissance. Soutenir la recherche dans ces domaines est donc primordial. Parmi les exemples d'applications et de bénéfices attendus, nous pouvons citer :

- l'aide au diagnostic, en particulier dans les cas difficiles, la planification thérapeutique ou chirurgicale, le pronostic, la recherche de biomarqueurs (dans des images médicales par exemple) ;

- la médecine interventionnelle, le geste thérapeutique ou chirurgical assisté par robotique;
- les dispositifs médicaux (prothèses, implants, délivrance de médicaments, e-santé, télémédecine, robotique,...) ;
- le résumé de données ;
- la prise en compte de plus de dimensions pour le raisonnement (données multi-sources, multi-échelles, dynamiques, interactions médicamenteuses,...) pour améliorer le service rendu au patient (un exemple est la consultation pré-anesthésie), et en contexte multi-patient en particulier pour les questions de santé publique et des analyses à grande échelle, ou encore la constitution d'atlas des maladies et des prescriptions ;
- la médecine des 5P (préventive, personnalisée, précise, participative, prédictive), s'appuyant en particulier sur le champ émergent des "omics" : ce champ promet un changement de paradigme et des avancées significatives, où ce ne sont plus les symptômes qui guideront principalement le diagnostic, mais où les patients seront aussi traités en fonction de leur propre génome, épigénome, transcriptome, protéome, métabolome, antibodyome, etc.[13];
- la modélisation et la simulation, d'une part pour comprendre un mécanisme lié à la santé, et d'autre part pour guider l'intégration de connaissances et l'analyse ou l'interprétation de données ;
- l'aide à la rédaction de comptes rendus médicaux, faisant appel à des techniques de TAL;
- l'aide à la codification des actes médicaux (PMSI) à partir du type de données ou d'examens et des comptes rendus ;
- l'extraction d'informations à partir de comptes rendus médicaux textuels pour l'utilisation secondaire des données de santé et l'aide à l'exploitation des connaissances publiées dans la littérature scientifique (pour alimenter des bases de données sur les résultats actuels de la science (les «machines à lire»)) : détermination des informations, qualification de leur degré de certitude, provenance de cette information et chaînes de citation, pour la recherche d'informations précises pour aider la sélection des articles scientifiques à des fins de méta-analyse, pour faciliter l'accès des praticiens aux données actuelles de la science en rendant aisée la recherche bibliographique dans leur langue, en leur permettant de poser des questions précises en langue naturelle, et en leur fournissant une traduction automatique des résumés des publications internationales (notamment les méta-analyses comme les revues Cochrane);
- l'aide à la publication des résultats de recherche par assistance à la rédaction non biaisée des parties-clé d'un article scientifique (affirmations / revendications dans le résumé et la conclusion);
- l'automatisation ou la semi-automatisation de tâches répétitives ou continues (par exemple pour le suivi d'une profondeur d'anesthésie) ;
- les dispositifs à destination du grand public pour le bien-être et la prévention, l'assistance à l'observance, le suivi thérapeutique (50% des médicaments prescrits ne sont pas pris ou mal pris (horaires, doses...));
- l'assistance à la personne (autonomie, handicap, réadaptation et rééducation, prédiction des chutes, vieillissement,...);
- l'explication des connaissances et de l'activité intellectuelle des médecins ;

[13] Par exemple pour la génomique, on peut citer le projet France Génomique 2025 : http ://social-sante.gouv.fr/systeme-de-sante-et-medico-social/recherche-etinnovation/france-genomique

- la formation des médecins et plus généralement des personnes impliquées dans les parcours de soin, notamment par la simulation à l'aide de patients virtuels et d'avatars numériques de vrais patients ;
- la lutte contre l'errance médicale, quel que soit le lieu d'habitat ou de travail (suggérer le bon diagnostic et orienter vers les bons spécialistes représente un enjeu majeur en termes de bonne prise en charge, de bien-être du patient, d'économie (élimination d'examens inutiles) et de qualité des soins, avec un enjeu en particulier pour les maladies rares) ;
- la formation thérapeutique des patients, notamment à l'aide de jeux sérieux et de mondes virtuels.

Verrous et défis

Pour répondre aux questions ci-dessus, quelques points constituent actuellement des verrous, qu'il faudra lever à court terme.
Le premier est certainement la question des **données**, en particulier à cause de la nature personnelle des données médicales et de santé, de la « « zone grise » sur ce que l'on peut considérer comme données de santé et de l'agrément indispensable pour stocker ces données. Si ce thème est largement mentionné depuis plusieurs années, il n'existe pas à ce jour de solution satisfaisante pour répondre aux besoins croissants d'accès, de partage, de protection, d'annotation, de garantie de fiabilité et d'absence de biais, en particulier pour la réutilisation de données dans des contextes différents de ceux où elles ont été collectées.
Des structures telles que le SNIIRAM pour les données de l'assurance maladie existent, mais restent encore trop limitées. Parmi ces données, celles d'imagerie constituent un volume très important et de grande complexité, ce qui augmente encore les difficultés de partage et d'annotation. Quelques initiatives larges voient le jour (Sage http://sagebase.org/, plateformeIAM dans le projet FranceLifeImaging...), mais la plupart des autres projets restent cantonnés à un petit périmètre d'équipes et d'applications. Cette question est en soit un problème d'IA, et par exemple l'annotation pourrait être accélérée et fiabilisée avec des méthodes de raisonnement associant ontologies et TAL.
De plus, l'intégration de données dans les systèmes hospitaliers nécessite des recherches sur la conception de ces systèmes et sur les interfaces, en fonction de l'opérateur et de l'utilisateur.
De même, la question de l'intégration des données patient dans un dossier patient unique reste posée, avec les problèmes questions associés d'interfaces et de remplissage. Des exemples dans les pays nordiques montrent que les dossiers patients uniques et centralisés sont un support pour la recherche par exemple en dépistage.
Pour les données textuelles, la possibilité de mise à disposition contrôlée aux équipes de recherche de données initialement privées est le frein principal dans l'avancée des travaux en France : la qualité des méthodes d'anonymisation doit être renforcée et des avancées faites sur des méthodes permettant de travailler sur des données chiffrées. La fiabilité des informations élémentaires extraites de textes est un autre enjeu : les extractions automatiques doivent qualifier ces informations (affirmées par l'auteur, niées, hypothétiques, conditionnelles, etc.), la fiabilité de leur provenance (autorité du support de publication pour la littérature scientifique), ainsi que la confiance que l'on peut accorder au processus d'extraction pour chaque information selon les moyens employés pour la déterminer. Quel que soit le type de données, leur robustesse et leur reproductibilité sont un souci majeur. Un certain nombre d'articles, souvent publiés dans les meilleurs journaux, ne semblent pas reproductibles et donc leurs conclusions prêtent à interrogation. Comment alors éviter une accélération du processus via une méthode d'IA qui intègrerait trop vite des données peu fiables ? Des programmes de recherche en sciences cognitives pourraient mieux analyser la relation du médecin aux données d'un système d'IA et comment les propositions du système influencent sa décision médicale.

Si les techniques d'apprentissage, où toute l'information est recherchée essentiellement dans les données, sont actuellement un domaine important, celles modélisant et intégrant les **connaissances**

sont également des défis : élicitation des connaissances, modélisation, combinaison avec des données individuelles ou de groupe, modèles de raisonnement et de décision, ingénierie ontologique et des connaissances... Par exemple, sur les aspects d'ontologies, la demande est forte dans de plus en plus de spécialités médicales ou périphériques aux spécialités (réseaux de soin/santé), avec une partie handicap (et sa modélisation) qui va inévitablement appeler de plus en plus de recherches (avec le vieillissement de la population et les maladies chroniques et neurodégénératives), dans un objectif de partage et d'augmentation de la connaissance. La disponibilité de ressources sémantiques en français (terminologies, ontologies) est un autre facteur limitant par rapport aux travaux d'analyse de textes en anglais.
Du point de vue de la recherche, les travaux à renforcer concernent l'aide à la décision, l'apprentissage de la sémantique et le traitement automatique des langues naturelles.

L'acceptabilité des techniques d'intelligence artificielle passe vraisemblablement par une meilleure interprétabilité des méthodes et des résultats qu'elles produisent. À titre d'exemple, un système évolutif qui apprendrait en continu au fur et à mesure de son utilisation est très difficile à faire accepter pour un marquage CE (Communauté Européenne) ou FDA (Food and Drug Administration), entre autres à cause de la difficulté de la définition de protocoles de marquage et à cause de la nature de la validation des performances cliniques qui ne doit pas être remise en cause par une évolution d'un système. Cette question montre le besoin de sensibilisation du monde de la normalisation à l'IA (et inversement celle du monde de l'IA aux questions de normalisation). Plus largement, l'acceptabilité concerne les cliniciens et praticiens, mais aussi les patients, les institutions, les entreprises, etc. Un axe important concerne le développement de modèles et d'algorithmes qui respectent par construction la préservation de la vie privée. Il s'agit alors de quantifier avec quel degré ces algorithmes respectent un modèle de préservation de la vie privée. Avec l'émergence de tels algorithmes, la confiance dans leur usage sera considérablement renforcée, et ainsi la mise à disposition de grandes masses de données sensibles (images, génomes, etc.) se trouvera facilitée. La dimension de l'usage doit également être prise en compte (avec les interfaces notamment) pour une acceptabilité et une utilisation adéquate. Le bénéfice clinique des systèmes d'IA mis en œuvre doit aussi être prouvé. Les recherches en informatique médicale évoluent dans ce sens et de nombreux laboratoires des usages se créent en santé, certains reprenant ou reconstruisant des habitats (sur)connectés, d'autres s'intéressant directement aux interfaces des systèmes informatiques (comme EVALAB à Lille). Les questions de visualisation des modèles et des données sont également à prendre en compte pour augmenter l'acceptabilité. Enfin, les questions de cybersécurité (par exemple pour les dispositifs médicaux) doivent être traitées car elles sont fondamentales dans le domaine de la santé.

La question de la **souveraineté** mérite une attention particulière, dans un contexte où les pratiques et les règles diffèrent considérablement d'un pays à l'autre. L'accès aux données étant une des clés de l'IA, faudrait-il donner les moyens aux entreprises françaises d'accéder en priorité aux données françaises ? La question de la réglementation des capteurs de données individuelles par les *smartphones* doit également être posée. De telles données peuvent avoir trait à la santé, même si elles ne sont pas considérées explicitement comme des données de santé (par exemple les données de localisation des personnes peuvent donner des informations sur la forme physique de l'individu, les établissements de soin visités, etc.). Enfin, la souveraineté peut se décliner, d'une part, en souveraineté nationale et, d'autre part, en souveraineté des individus sur leurs données (en lien avec le point suivant).

Les questions de **réglementation et d'éthique** sont évidemment particulièrement sensibles dans le domaine de la santé et de la médecine, avec celle de l'agrément, et évoluent rapidement (loi Jardet récente par exemple). Nous distinguons ici les questions éthiques, comportant discussion ouverte, mise en évidence des tensions, débats sur la stratégie considérée comme la meilleure à un instant donné, et les questions réglementaires qui correspondent à la mise en oeuvre normative éventuelle des stratégies éthiques et à l'établissement d'un code de déontologie du domaine qui, dès lors qu'il existe, ne prête plus à une discussion ouverte mais appelle à un respect sanctionné en cas d'infraction. De nombreux

travaux traitent du sujet, d'une part dans le domaine de la recherche et, d'autre part, concernant la prise en charge des patients, les données, les règles de remboursement, la protection de la vie privée, etc. Les principes éthiques (autonomie de la personne, bienveillance/non-malveillance, justice) doivent être préservés et donc les moyens techniques adaptés à cette préservation. Par exemple l'autonomie implique le consentement libre et éclairé. Il faut donc inventer les moyens d'exercice de ces caractères libre et éclairé et donc trouver les moyens d'une information continue des personnes permettant un éventuel droit de retrait ou une interpellation si l'usage des données est perverti ou frauduleux. L'IA peut elle-même probablement venir au secours de la préservation de ces droits car seuls des systèmes d'analyse des données massives pourront à l'avenir repérer des usages inappropriés des données des patients. Compte-tenu de la rapidité d'évolution du domaine, des incertitudes importantes et des interrogations multiples, l'organisation d'un processus de réflexion continu impliquant les différents acteurs, spécialistes et non-spécialistes, est essentielle. Ce processus doit inclure le débat public mais ne peut se limiter au débat public. Il implique des instances de réflexion comme la CERNA, qui travaille sur plusieurs sujets sur l'éthique dans la recherche en numérique, dont la robotique et les données (http://cerna-ethics-allistene.org/), et pourrait donner lieu à un rendez-vous régulier de type "Etats généraux du numérique et de l'IA" qui permettraient de débattre sur la place publique des questions d'usage et de règlementation.

Actuellement les possibilités de **transfert** des travaux de recherche vers l'industrie sont considérablement freinés par de multiples aspects : accès aux données, réglementations, difficultés de financement, disponibilité des chercheurs... Citons quelques exemples à l'étranger de collaborations lancées entre des industriels et des équipes de recherche en IA et imagerie : le MGH a mis 25M$ dans un *Clinical Data Science* dédié à l'IA pour la radiologie, également soutenu par Nvidia, le lancement d'un *AI research lab* à Montreal, avec plus de 100M$ autour de Yoshua Bengio, qui est aussi fondateur d'Imagia, une startup dédiée à la problématique imagerie et IA, le *Center for digital health innovation* lancé récemment entre UCSF et GE Medical pour travailler spécifiquement sur les problématiques d'imagerie médicale, le *Data institute* de USFCA piloté par Jeremy Howard, qui a remporté de nombreux challenges Kaggle et a créé Enlitic, la première startup dédiée aux problématique IA et imagerie médicale. En France, les projets sont en retrait par rapport au potentiel, alors que c'est un enjeu économique très important. Il convient également de mentionner de très gros investissements aux Etats-Unis pour créer des infrastructures pour les données de santé : HITECH, avec 18 Mds$ pour équiper hôpitaux et cliniques avec des EHR (*Electronic Health Record*) et 2 Mds$ pour des systèmes d'échanges de données, NIH avec 616M$ pour le projet Big Data Knowledge et 216 M$ pour le projet Precision Initiative Medecine, etc.

Recommandations

Les recommandations générales des groupes Recherche amont et Transfert s'appliquent dans le domaine de la santé et de la médecine. Soulignons-en quelques-unes qui ont des impacts sur les aspects de transfert :

Accès aux données de santé et éthique : donner les moyens à la recherche, au développement et à la création de produits de travailler avec des données de santé tout en respectant leur confidentialité.

- Développement et renforcement de l'approche française opendata (data.gouv.fr), et favorisant la constitution de bases de données médicales et de santé annotées, ainsi que de méthodes et d'algorithmes ouverts. L'initiative européenne ``free flow of data'' devrait être examinée et discutée, en particulier au regard des entités et entreprises qui l'ont poussée.
- Engagement d'une démarche de simplification de la réglementation encadrant l'usage des données de santé pour la recherche ainsi que la dé-identification et la ré-identification des données, tout en veillant à garantir l'éthique de l'usage et la sécurité des données. Les instances réglementaires et administratives doivent également avoir un rôle de conseil du bon usage de la réglementation et des outils associés.

- Donner les moyens à la CNIL pour qu'elle puisse donner des avis plus rapide pour permettre la mise en place, dans de bonnes conditions et avec de bonnes garanties sur la protection des données personnelles, de flux de données médicales accessibles aux chercheurs et aux entreprises.
- Collecter ou créer et mettre à disposition un éventail de solutions techniques permettant de faire de la recherche et du développement sur des données de santé en préservant leur confidentialité, et créer un guide des bonnes pratiques pour le choix et l'usage de ces solutions.

Intégration IA et Santé : développer les compétences en intelligence artificielle avec et dans le monde de la santé.
- Soutien à la recherche multidisciplinaire dans ce domaine, afin de pouvoir répondre à l'ensemble des questions soulevées dans ce rapport.
- Chaires mixtes dans les hôpitaux, éventuellement associées aux structures des IHU et RHU.
- Formation aux nouveaux métiers et à leur évolution avec le développement du numérique et de l'IA, et de personnes maîtrisant l'ensemble des techniques de l'IA et capables de décider laquelle ou lesquelles mettre en œuvre. Dans le même ordre d'idées, actualisation des formations de santé (initiales et continues) aux nouveaux outils, pour les médecins, mais aussi pour les autres personnels de santé.
- Encouragement des carrières public / privé.

Financements et partenariats.
- Mise en place d'un fonds d'investissement dédié aux technologies d'intelligence artificielle en santé.
- Encouragement des partenariats entre partenaires publics privés (financement généralisé sur le modèle des IRT).

Renforcement et développement des recherches et des outils dans le domaine du numérique pour une meilleure prise en charge des patients.
- Soutien au développement de la médecine préventive, et plus généralement des 5P.
- Développement de la télémédecine, de la notion de patient virtuel numérique, dont un bénéfice direct est la numérisation de l'ensemble de la prise en charge du patient.
- Mise en place du DMP.

2.2.2.4 Energie

2.2.2.4.1 Contexte

Les énergies renouvelables sont aujourd'hui arrivées à parité réseau et le développement puis l'adoption par le monde industriel de l'intelligence artificielle font de l'IA un élément clé de la quatrième révolution industrielle. Résultat : une mutation profonde de l'écosystème énergétique est en cours :

- Le renouvelable, compétitif, plus ou moins distribué, redéfinit progressivement les cartes du jeu électrique. Un sujet sensible et stratégique.
- La digitalisation du réseau "smart grid" semble ouvrir de nombreux gisements de valeur depuis l'échelle hyper-locale jusqu'à l'échelle nationale et européenne.
- La tendance est à l'interconnexion massive et les villes pourront demain bénéficier d'un véritable Internet de l'Énergie.

En France, le système existant est encore très centralisé, conçu pour assurer un accès égalitaire à l'électricité, encourager une compétitivité industrielle, et une certaine indépendance énergétique basée sur le nucléaire, porté en particulier par le principe de solidarité et la péréquation.

L'IA un acteur clé de la mutation de l'écosystème énergétique

L'Intelligence artificielle, qui a pour but de permettre à un ordinateur d'exécuter une tâche complexe, par apprentissage des données observées ou grâce à un modèle formel, représente un levier important pour faire face à la complexité des nouveaux problèmes rencontrés, notamment :

- optimiser les processus et réaliser des économies d'énergie (ex: maintenance prédictive, optimisation du plan de production, adaptation en temps réel des paramètres opératoires, etc.)
- planifier et piloter la transition énergétique grâce à des décisions guidées par la donnée (ex: simulation fiable des problèmes d'intermittence sur le réseau lorsque la pénétration des ENR augmente).

Problème 1: Efficacité et flexibilité de la demande

Dans un système électrique local avec un fort taux de pénétration de sources de production intermittentes, la valeur de prévision de l'offre et d'adaptation de la demande électrique en fonction de la production disponible, est démultipliée. Néanmoins, l'usage, l'expérience (quelle qu'elle soit, industrielle, commerciale, mobile, résidentielle, via des objets connectés, etc.) ne doivent pas être impactés, pour permettre et garantir une adoption massive.

De plus, les nouveaux usages qui se développent (notamment la mobilité, les bâtiments intelligents, les "prosumers") apportent leur lot de contraintes bien sûr mais donnent aussi plus de flexibilité au système.

Problème : Comment révéler et activer l'efficacité et la flexibilité de la demande, sans impacter l'usage, voire en l'améliorant?

C'est une des priorités de la transition énergétique aujourd'hui. Et de même que l'on demande aux citoyens de recycler leurs déchets, on peut les inciter à participer à la flexibilité pour des raisons qui ne sont pas uniquement économiques. Il y a un vrai gisement de "valeur sociale" dans la production collective du "bien commun".

L'IA comme élément de solution :

Il existe de nombreuses efficacités et flexibilités à exploiter. La première étape consiste à sélectionner les données ciblées, abondantes (big data) d'usage croisés, typiquement :

- Résidentiel: Chauffage, Machine à laver, Froid, Eclairage…
- Ville: mobilité, sécurité, services municipaux prioritaires…
- Industrie: demande client, fiabilité, logistique…

L'IA peut permettre de dégager des profils de consommation efficace et flexible voire d'activer certaines consommations, sans "dégrader" l'usage :

- SmartHome et Bâtiments intelligents : améliorer la gestion énergétique des bâtiments grâce à l'IA (pilotage, effacement, optimisation des consommations, automatisation GTB)
- Eco-mobilité et optimisation des réseaux de transports : faciliter les déplacements et limiter la congestion urbaine (bus ou voitures électriques, gestion des bornes de recharge et de

l'utilisation du trop-plein des batteries électriques, gestion croisées des systèmes multi fluides, gaz, chaleur, hydrogène)

L'IA peut promouvoir la recherche de solutions décentralisée et coopérative d'un équilibre entre les besoins consommateurs et la nécessité de stabilité du réseau électrique.

Problème 2: Équilibrage fin du réseau

Le développement de sources renouvelables distribuées, intermittentes, exige de travailler à des échelles plus locales. A cette maille, le bruit des productions intermittentes et des consommations plus discrètes et spécifiques demande une capacité de prévision et d'équilibrage infiniment plus fine, en maille spatiale proche de l'objet connecté, et en maille temporelle, proche du temps réel.
En remontant récursivement en "poupées russes", ces équilibrages fins servent le réseau à l'échelle nationale voire européenne.

Problème : Comment assurer cet équilibrage à la maille fine?

L'IA comme élément de solution :

La prévision fine de la production intermittente est critique pour l'équilibrage. Application naturelle pour l'intelligence artificielle, basée sur l'intégration de données externes ciblées (par exemple prévision météo – ensoleillement, force du vent, probabilité de passage d'un nuage très localement), maillées géographiquement et dans le temps.

La prévision fine de la consommation devient tout aussi primordiale. Probablement plus complexe que la production, elle se base sur l'usage (voir Problème 1). L'enjeu est de révéler la flexibilité pour épouser la production intermittente. Ceci demande donc d'activer des mécanismes d'adaptation de la demande, ciblé dans l'espace et le temps.
L'IA permet une gestion énergétique automatisée en recherche d'équilibre offre-demande analysée en temps réel et coordonnée des réseaux d'électricité, de gaz, et de chaleur. Cette gestion à la maille de la ville, locale, se traite de manière récursive pour remonter à grande échelle l'ensemble des réseaux intelligents pour avoir une gestion globale de l'énergie, basée sur des briques locales.

Il serait donc envisageable d'obtenir un système avec de nombreuses IA, maille par maille, néanmoins, la question se pose d'une IA souveraine, ou bien d'un système de coordination souverain, capable d'assurer l'équilibrage du réseau, de permettre l'équilibrage à maille plus grosse, jusqu'au système national et européen.

L'IA ouvre des perspectives très prometteuses pour permettre, de manière récursive, l'équilibrage depuis les mailles les plus fines, jusqu'au système national.

Problème 3: Design de marché et réglementation

La structure de marché, réglementée, reflète une prise de position politique. Assurer l'accès équitable à tous de l'électricité, l'accommodation de renouvelables en majorité, le choix des acteurs d'échanger de l'énergie autoproduite, voire des services au réseau, la taxation sur la valeur créée, sont autant de sous-jacents au prix et aux mécanismes de marché. Aujourd'hui la gestion (sur des mécanismes type marché d'ajustement) se fait à l'échelle du réseau national. Il y a un enjeu à développer des mécanismes au niveau de la boucle locale.

Problème : Comment structurer un marché de l'électricité qui reflète la stratégie nationale et européenne?

L'IA comme élément de solution :

En prenant en compte les contraintes décidées souverainement, l'intelligence artificielle peut aider au design de structure de marchés, plus subtil, plus local, en s'inspirant de l'existant et de marchés dans des filières adjacentes.

Un autre avantage de l'IA est de proposer un nouveau support (et non substitution) décentralisé à la décision. En effet, pour débloquer les postures partisanes respectives des acteurs, l'apport de préconisations documentées et tracées, mandatées par une structure souveraine, pourrait aider à débloquer les discussions, en fonction des décisions politiques.

En creux, l'action proactive du gouvernement français sur le sujet permettra de conserver le contrôle et de générer la valeur localement.

Problème 4: Visibilité de la donnée

Pas de surprise, l'Intelligence artificielle s'alimente de données.
Si la digitalisation du réseau est nécessaire, et en cours, cela n'est pas suffisant. Il faut rendre visible la donnée du réseau électrique à la filière pour pouvoir maximiser la création de valeur depuis la maille ultra-locale jusqu'à la maille nationale, sur de nombreuses échelles de temps. Topologie fine de l'existant, profils de consommation, profils de production, historiques de données, historique des incidents réseaux, entre autres.
Cette donnée directement captée, ou bien calculée, va être progressivement prévisible et accessible aux acteurs qui concentrent la puissance de calcul. Typiquement les GAFA. Impliquant un enjeu national de compétitivité, de cybersécurité.

Problème : Comment promouvoir la filière française grâce à un accès contrôlé aux données?

L'IA comme élément de solution :

La visibilité de la donnée devient un choix stratégique, donc politique et souverain.
Quelle donnée est jugée critique? Quelle donnée est jugée publique? Quelle valeur a quel acteur. La question n'est pas nouvelle.

Néanmoins, il semble critique, du point de vue de l'énergie, d'adresser en plein ce sujet. Le groupe recommande de proposer une approche de visibilité. Tel jeu de données est visible pour tels acteurs. Cela permet de gérer plus finement la valeur. Ainsi que promouvoir la standardisation des formats d'échange et des moyennes d'échange de données afin de faciliter l'interopérabilité entre les acteurs du secteur (producteurs, prosumer, consommateurs, collectivités territoriales, équipementiers). Assurer l'accès équitable et règlementé aux données.

L'IA implique des croisements nouveaux de jeux de données. La visibilité, contrôlée, est critique pour en extraire la valeur.

2.2.2.4.2 Recommandations

Nous proposons deux recommandations “pré-requis”:

1. Visibilité
2. Avatar

Puis deux recommandations “challenges”:

3. Equilibrage
4. Marché

- **Prérequis 1: Visibilité de la donnée**

L’objectif est de comprendre l’ensemble des données existantes liées à l’énergie, identifier leur sensibilité, leur capacité à être déduites.
Le livrable est de proposer une matrice de visibilité, pour débloquer le cloisonnement et se donner une chance de concourir pour la création de valeur de l’IA, nationalement.

- **Prérequis 2: Avatar du réseau réel**

L’objectif est de se doter d’un moyen de simulation reproduisant la structure, le fonctionnement et la dynamique du réseau électrique réel. Notons que EDF R&D et IBM ont déjà réalisé des travaux de recherche sur cette thématique [1].

Cet “avatar” représentant le réseau électrique réel permettra d’évaluer l’intérêt des approches d’IA *(avant de les appliquer sur les systèmes réels)*. D’abord, le réseau électrique peut être modélisé par un “graphe” où chaque nœud représente un équipement *(ex: compteur, transformateur, poste source, etc.)*. Puis l’idée est de simuler la consommation et la production d'électricité en chacun des nœuds du réseau.

Les approches d’IA permettront de simuler cela au niveau le plus fin *(i.e. les compteurs individuels)* et d’agréger les consommations en remontant la hiérarchie du réseau. Ensuite, ce modèle nous permettra de simuler des scénarios de changements des nouveaux usages et des moyens de production ENR.

- **Challenge 1: Equilibrage**

Avec la libéralisation progressive du marché et la révolution des renouvelables, une valeur locale significative peut être générée autour de la stratégie d’équilibrage locale. Ces équilibrages peuvent se consolider à la manière de poupées russes et la valeur transférée ainsi jusqu’à l’échelle nationale.

Nous recommandons l’organisation d’un challenge de solution d’équilibrage, sur la maille du choix du candidat, basé sur des solutions IA.

Avec des contraintes de l’infrastructure, les candidats proposeront des nouveaux mécanismes d’équilibrage, pour favoriser un objectif précis. 100% Renouvelable, fiabilité, minimum d’upgrade d’infrastructure, solidarité, génération de valeur locale…

- **Challenge 2: Marché**

Avec la libéralisation progressive du marché, et la révolution des renouvelables, une valeur locale significative peut être générée autour de la conception de places de marché locales.

Nous recommandons un grand concours de conception de places de marché, sur la maille du choix du candidat, basé sur des solutions IA.

Avec des contraintes souveraines de péréquation et prélèvement de la valeur, les candidats proposeront des nouveaux mécanismes de marché, pour favoriser un objectif politique précis. 100% Renouvelable, fiabilité, solidarité, génération de valeur locale…

Ethique & Cybersécurité[14]

Les réseaux électriques à toutes les échelles deviennent "smart" et deviennent également des réseaux dans lesquels de l'information est récupérée, circule, est échangée, interprétée, etc. entre les différents points, qu'ils soient petits ou gros consommateurs, petits ou grands producteurs, flexibles ou non, prévisibles ou non, voire des nœuds du réseau ou des objets connectés adjacents. Cette complexité accrue apporte de nouveaux challenges en termes de sûreté, combinant sûreté du réseau (et la sécurité d'approvisionnement correspondante), sûreté de l'information (informations personnelles, données process sur des consommateurs essentiels, données business, etc.), sûreté des installations de production ou de consommation. Les scénarios catastrophes sont nombreux, et pour certains réels, comme par exemple la "prise en otage" de l'approvisionnement électrique (exemple de l'attaque sur les réseaux de distribution ukrainiens en 2015)

L'IA a un rôle à jouer à différents niveaux

- Conception des systèmes et des réseaux (capacité à simuler et à anticiper les points faibles)
- Détections de signaux faibles précurseurs d'attaque
- Capacité à transférer rapidement d'une boucle du réseau à une autre les apprentissages pour détecter les menaces émergentes (en adaptant les méthodes de détection à la topologie locale du réseau)

2.2.2.4.3 Exemples de projets

Exemples de projets industriels

Prévision de consommation électrique et de facture individuelle à la fin du mois
En utilisant les données historiques de consommation d'un site (compteur d'électricité général) ainsi que des données caractéristiques du site (surface, étage, occupation, ...) et des données météorologiques, l'IA élabore des modèles permettant de prédire les volumes des différents types d'heures du contrat de ce site (heures pleines / creuses / pointes) ainsi que la prévision des factures. La sortie du modèle est le volume total de consommation du mois. La précision de cette estimation s'améliore au fur et à mesure que l'on obtient plus d'informations au cours du mois étudié.

Désagrégation de courbe de charge avec ou sans capteurs.
En partant d'un ensemble réduit de sites équipés d'une instrumentation précise qui mesurent la consommation électrique de chacun des usages (chauffage, climatisation, éclairage, serveurs, …), il est possible d'entraîner à l'aide du Machine Learning un modèle capable de généraliser à un plus grand nombre de sites similaires. Le modèle prend en entrée la consommation instantanée mesurée au compteur général, ainsi que des données météorologiques et des caractéristiques du bâtiment et prédit en sortie la consommation de chacun des usages sur le site.
Il est possible de grandement améliorer la performance de ce type de modèle en réalisant une installation temporaire de capteurs sur les sites étudiés (pour une durée minimum d'une semaine), pendant une période prédéfinie de l'année. Cela fournit au modèle une information sur la répartition normale des usages sur le site et améliore la précision des prévisions, même sur des périodes très différentes de celle où avait été posés les capteurs.

[14] https://www.ifri.org/sites/default/files/atoms/files/desarnaud_cybersecurite_2017_sl.pdf
http://www.smartgrids-cre.fr/index.php?p=gestion-donnees-cyber-securite
http://data.over-blog-kiwi.com/1/87/83/64/20151031/ob_b39ecc_livre-blanc-sur-la-cybersecurite-des-r.pdf

Analyse comportementale et conseil en temps réel d'optimisation de la consommation énergétique
Les consommations d'un bâtiment (par usage ou mesurées au compteur général) et ses caractéristiques, ainsi que des données comportementales ou techniques (type d'éclairage, température cible de chauffage, comportement vis-à-vis de l'éclairage) permettent de construire un modèle « génératif » du bâtiment. En utilisant celui-ci, il est possible de quantifier l'impact d'un changement de comportement sur les consommations énergétiques d'un bâtiment. L'IA permet aussi de fournir des conseils en temps réel afin d'aider à l'identification de postes consommateurs et par analyse de la courbe de charge optimiser le comportement dans le but d'optimiser les consommations.

Maintenance prédictive
L'utilisation de données provenant de SIG (systèmes d'informations géographiques) tels que le positionnement de câbles à haute tension, leurs dispositions géographiques, ainsi que des données comme leur longueur ou leur année de construction permet d'entraîner à l'aide du Machine Learning un modèle capable de donner la probabilité qu'une panne ou un claquage arrive sur une période future donnée.

Exemples de projets institutionnels

Bâtiments publics intelligents
Dans le but d'optimiser la consommation d'énergie des infrastructures, l'IA apprend sur base de l'historique de données collectées par des capteurs installés sur les équipements du bâtiment et est capable, sur la base de Machine Learning ou Deep Learning, de baisser la consommation énergétique du bâtiment soit en fournissant les bons conseils aux techniciens qui pilotent les usages, soit l'IA pilote directement la GTB du bâtiment.

SmartGrids Vendée

Dans le cadre des Smartgrids, l'intelligence artificielle en se basant sur les données historiques ou temps réels des moyens de production (classiques et renouvelables) mais aussi de consommations à l'aide des compteurs intelligents et des capteurs au sein des villes intelligentes, permet d'élaborer l'ensemble des modèles de prédictions permettant une gestion optimisée du réseau ainsi que des usages des consommateurs finaux.
La gestion des contraintes de distribution prend une autre dimension avec l'arrivée massive des énergies renouvelables (intermittence) et l'apport de l'intelligence artificielle au cœur de l'ensemble des processus de gestion et de pilotage du réseau permet d'optimiser l'offre et la demande d'électricité.
En plus des modes de productions renouvelables, viennent s'ajouter les nouveaux modes de stockage de l'énergie (stockage hydrogène, batterie des véhicules électriques, batterie en maison) qui constituent là aussi la nécessité d'élaborer des algorithmes d'IA capable de gérer les contraintes de déphasage et d'évaluation des impacts sur l'ensemble du réseau (équilibrage réseau, maintenance prédictive des câbles).
L'IA en agrégeant l'ensemble des informations sur les ressources disponibles et demandées, permet d'éviter tout dysfonctionnement du réseau et procure à l'aide de nouveaux outils numériques un conseil en temps réel au l'usager afin de lui permettra à la fois de maîtriser ses consommations mais aussi de bénéficier d'une meilleure qualité de service.

Exemples de projets de recherche

Prédiction de l'erreur de zone de contrôle et équilibrage réseau
Les gestionnaires de réseau de transports d'électricité peuvent bénéficier d'un modèle capable de prédire les erreurs d'équilibre entre génération et demande. Un tel modèle utilise les données historiques, ainsi que les programmes d'échanges planifiés dans le futur et les prévisions de

production et de charge pour prédire à courte échéance (30 minutes) avec une granularité fine (30 secondes) les erreurs de zone de contrôle. Un tel modèle pourrait être intégré à un module de contrôle de fréquence (LFC) dans le cadre de l'activation d'un système de réserves de restauration de fréquence (aFRR)

Mobilité électrique

Si tous les véhicules électriques en France se chargeaient tous en même temps, ils pourraient consommer la moitié de la puissance installée en France.

Le machine Learning permet d'élaborer des modèles de prédictions à partir des courbes de charge et décharge, ainsi que de la puissance des véhicules et de l'historique d'utilisation, afin d'optimiser la charge du véhicule électrique sans mettre en risque le réseau.

Pour une entreprise, ces modèles peuvent aussi, en cas de surcharge réseau, permettre d'utiliser les batteries de véhicules pour alimenter ses propres équipements.

Prévision de courbe de charge à la maille locale

A partir des données de météo, de couverture nuageuse, mais aussi d'activité économique, il est possible de prévoir les profils de consommation à la maille nationale jusqu'à la maille territoriale. Avec des algorithmes de Machine Learning, il est possible d'affiner la prévision jusqu'à la maille locale voir d'un quartier

Références bibliographiques

[1] Pompey, P., Bondu, A., Goude, Y., & Sinn, M. (2015). Massive-scale simulation of electrical load in smart grids using generalized additive models. In *Modeling and Stochastic Learning for Forecasting in High Dimensions* (pp. 193-212). Springer International Publishing.

3. Annexes

3.1. Composition du groupe et des sous-groupes

Les membres du groupe principal ont été :

Académiques
Bertrand Braunschweig, Inria (co-animateur)
David Sadek, IMT (co-animateur)
Olivier Boissier, Ecole des Mines de Saint-Etienne
Ludovic Denoyer, Université Pierre et Marie Curie – LIP6
Luis Farinas del Cerro, Institut de Recherche en Informatique de Toulouse
Catherine Faron Zucker, Université de Nice – I3S et Inria
Jean-Gabriel Ganascia, Université Pierre et Marie Curie – LIP6
Christine Solnon, Laboratoire d'InforRmatique en Images et Systèmes d'information
Philippe Watteau, CEA-LIST assisté de *Julien Chiaroni,* CEA-LIST
Isabelle Bloch, IMT, animatrice du sous- groupe Santé
Gilles Dowek, Inria, animateur du sous- groupe eEducation
Arnaud de La Fortelle, IMT, animateur du sous- groupe Robotique

Entreprises
Patrick Albert, SuccessionWeb (co-animateur)
Bastien Albizzati, Critéo
Caroline Chopinaud, craft ai, assistée de *Clodéric Mars*, craft ai
Stéphane Durand, Dassault Aviation, assisté de *Françoise Briand*, Dassault Aviation
Patrick Johnson, Dassault Systèmes, assisté de *Manuel Gruson*, Dassault Systèmes
Nadine Leclair, Renault, assistée de *Jean-Marc David*,
Juliette Mattioli, Thales
Pierre Riou, Acimeo, animateur du sous-groupe Sociétés de conseil et intégrateurs

Représentants des services de l'Etat
Medhi Gmar, Commissariat Général à l'Investissement
Cédric Nozet, Direction Générale des Entreprises
Tania Lasisz, SGMAP/Direction Interministérielle du numérique et du SI de l'Etat

Ainsi que
Stéphane Jourdain, Laboratoire National d'Essais
Juliette Kahn, Laboratoire National d'Essais

Sous-groupes thématiques

Cinq sous-groupes thématiques ont été constitués sur la robotique, la santé, l'e-Education, l'énergie et les prestations de conseil aux entreprises. La mission de chaque sous-groupe était de produire un court document sur les enjeux de l'industrialisation des recherches en IA pour son domaine.

e-Education
Gilles Dowek, Inria (animateur)
Pierre Collet, Icube, Université de Strasbourg
Amélie Cordier, LIRIS
Gabriel Cordier, Neuralcat
Jean-Marc Hasenfratz, Inria Learning Lab
Laurent Janin, Educlever
Alain Mille, LIRIS
Fabrice Moizan, Gaya Technologies
Catherine Mongenet, France Université Numérique (FUN)

Robotique
Arnaud de la Fortelle, IMT (animateur)
Philippe Bonnifait, Heudyasic
Guy Caverot, BA Systèmes
François Charpillet, Inria
Rodolphe Gélin, Softbank Robotics
Jean-Paul Laumond, LAAS

Santé
Isabelle Bloch, LTCI, Télécom ParisTech (animatrice)
Stéphane Buffat - IRBA, SSA-CNRS-Paris Descartes
Jean Charlet, LIMICS, AP-HP, INSERM, Université Pierre et Marie Curie
Hervé Chneiweiss - INSERM
Olivier Clatz, Therapixel
Alexandre Gramfort, LTCI, Télécom ParisTech & Inria
Aymeric Perchant - SATT Paris-Saclay
Gaël Varoquaux, Inria
Pierre Zweigenbaum - LIMSI, CNRS, Université Paris-Saclay

Energie
Yael Azoulay (co-animatrice), Air Liquide
Thomas Carrier (co-animateur), Air Liquide
Vincent Schachter, eMotorWerks & Engie
Alexandre Raulot, Expert Energie
Vincent Danos, Chercheur
Elie Rotenberg, Chercheur
Henri Binsztok, Chercheur
Tuncay Isik, Prevision.io

Bruno Aidan, Air Liquide
Alexis Bondu, EdgeML et CASA
Vincent Gauthier, Institut Mines-Telecom

Sociétés de conseil et Intégrateurs
Pierre Riou, Acimeo (animateur)
Patrick Albert, SuccessionWeb
Issam Ibnouhsein, Quantmetry
Sandrine Godefroy, Gregor Maciak, Thierry Ragin, Frédéric Bellaiche, Econocom
Frédéric Oblé, Wordline
Gabriel Kepeklian, Atos
Eric Monchalin, Atos
Didier Bourigault, Synomia
Maurice N'Diaye, Synomia
François-Régie Chaumartin, Proxem

3.2. Questionnaires

3.2.1. Questionnaire Recherche

Questionnaire « Industrialiser et transférer les technologies de l'IA »
Chercheurs

Cher(e) collègue

Dans le cadre du groupe de travail qui prépare un ambitieux Plan National pour l'Intelligence artificielle pour le compte du Secrétariat d'Etat au Numérique (Axelle Lemaire), et en collaboration avec le Secrétariat d'État chargé de l'Enseignement Supérieur et de la Recherche (Thierry Mandon), nous recueillons auprès de la communauté les informations relatives aux résultats ou travaux de recherche en Intelligence artificielle susceptibles d'être exploités d'ici trois ans ou moins, afin de définir des dispositifs de soutien spécifiques avec les moyens associés.

Notre travail s'étalera sur deux mois, mais il s'agit dans un premier temps d'identifier les potentialités de transferts, leurs modalités et les moyens à mettre en œuvre.

Merci de nous retourner ce questionnaire rempli avant le 10 février.

- - - - - - - - - - - - - - - - - - -

Identifiez les résultats relevant du domaine de l'Intelligence artificielle dont vous disposez, susceptibles d'être industrialisés et transférés pour une utilisation opérationnelle à court terme (moins de trois ans).

1. Indiquez les sous-domaines de l'IA dont relève votre recherche
2. Pensez-vous que ces sous-domaines soient suffisamment matures pour devenir une réalité industrielle à court terme ? Si non, pourquoi ?
3. Si oui, quel(s) est (sont) les secteurs « marché » adressés par ces sous-domaines ?
4. Disposez-vous de résultats que vous considérez transférables à court terme
5. Pour chaque résultat
   - Décrivez brièvement le résultat du point de vue scientifique et technique
   - Décrivez le problème concret qu'il résout et les solutions qu'il apporte, donner des cas d'usage
   - Qualifiez sur une échelle de 1 à 3 (1 : démonstrateur; 2 : preuve de concept ; 3 : système testé en conditions d'utilisation réelle) le degré de maturité technologique actuel de ce résultat
   - Décrivez succinctement les moyens requis pour l'industrialisation et le déploiement, et les étapes pour ce faire.
6. Avez-vous des expériences passées de transfert, positives ou négatives ?
7. Autres suggestions, recommandations

3.2.2. Questionnaire Entreprise

Questionnaire « Industrialiser et transférer les technologies de l'IA »
Entreprises

Cher(e) collègue

Dans le cadre du groupe de travail qui prépare un ambitieux Plan National pour l'Intelligence artificielle pour le compte du Secrétariat d'Etat au Numérique (Axelle Lemaire), et en collaboration avec le Secrétariat d'État chargé de l'Enseignement Supérieur et de la Recherche (Thierry Mandon), nous recueillons les besoins des entreprises en matière de technologies d'intelligence artificielle. Le but est de publier ensuite un ou plusieurs appels à projets en réponse à ces demandes.

Merci de nous retourner ce questionnaire rempli avant le 10 février (une fiche pour chaque besoin), ou sinon une réponse ouverte décrivant les opportunités que vous identifiez. Vous pouvez également nous indiquer si vous souhaitez interagir avec nous au-delà de ce questionnaire.

- - - - - - - - - - - - - - - - - - -

Questionnaire « Industrialiser et transférer les technologies de l'IA »

8. Quels cas d'usage pourraient bénéficier de l'intelligence artificielle dans votre entreprise à court terme (moins de trois ans) ?
9. Pour chaque cas d'usage:
   - Décrivez brièvement (une dizaine de lignes maximum) ce(s) besoin(s), le(s) problème(s) concret(s) à résoudre et les solutions à apporter ?
   - Quels avantages apporterait une solution intégrant de l'IA ? Pourriez-vous qualifier sur une échelle de 1 à 3, l'importance du bénéfice envisagé (1 : marginal, 2 : important, 3 : critique),
   - Que mettriez-vous en œuvre pour traiter ce cas d'usage (quelques exemples non exclusifs : recruter ou former, monter un partenariat avec un laboratoire de recherche ou un acteur privé, acquérir une technologie, etc.)
   - Quelles seraient les prochaînes étapes à franchir (maturation, expérimentation, industrialisation et passage à l'échelle, certification, déploiement) ?
   - Quels sont les freins éventuels à cette mise en œuvre ?
   - A quelle technologie d'IA pensez-vous pour traiter ce cas d'usage ?
   - Avez-vous identifié des compétiteurs nationaux ou hors de nos frontières qui ont déjà une mise en œuvre de type IA sur le sujet ?
10. Avez-vous des expériences passées de transfert, positives ou négatives ?
11. Autres suggestions, recommandations

3.2.3. Questionnaire Intégrateur

Questionnaire « Industrialiser et transférer les technologies de l'IA »
Intégrateurs

Dans le cadre du groupe de travail qui prépare un ambitieux Plan National pour l'Intelligence artificielle pour le compte du Secrétariat d'Etat au Numérique (Axelle Lemaire), et en collaboration avec le Secrétariat d'État chargé de l'Enseignement Supérieur et de la Recherche (Thierry Mandon), nous recueillons les besoins des entreprises en matière de technologies d'intelligence artificielle, afin de définir des dispositifs de soutien spécifiques avec les moyens associés.

Merci de bien vouloir renseigner ce questionnaire avant le 17 février, ou sinon nous adresser une réponse ouverte décrivant les opportunités que vous identifiez dans le champ libre en fin de formulaire (Question 4). Vous pouvez également nous contacter si vous souhaitez interagir avec nous au-delà de ce questionnaire.

-

0- Proposez-vous ou envisagez-vous à court terme de proposer des services, solutions, produits, plateformes s'appuyant sur des technologies d'IA à vos clients ?
1- Si oui, quels sont les 3 principaux services/solutions concernés
    a. Service n°x, x=1,2,3
        i. Décrivez brièvement le service proposé (10 lignes)
        ii. Quels avantages apporteraient une solution intégrant de l'IA ?
        iii. Sur quelles technologies d'IA s'appuie (rait) cette solution :
            1. A définir
            2. En cours de développement interne ou externe
                a. Si externe, labo ou start-up, autre
            3. Plateforme du marché (si oui, laquelle)
            4. Solution intégrée du marché (si oui, laquelle)
    b. Quelle sont les références (nombre) dont vous disposez pour ce service et le statut des références :
        i. En production, en déploiement, en prototype, en approche commerciale
    c. Quelles seraient, selon vous, les prochaînes étapes à franchir ? (ex. maturation, expérimentation, industrialisation et passage à l'échelle, certification, déploiement, etc)
    d. Quels sont les freins éventuels à cette mise en œuvre ?
    e. Avez-vous identifié des compétiteurs nationaux ou hors de nos frontières qui ont déjà une mise en œuvre de type IA sur le sujet ?
2- Avez-vous des expériences passées de transfert, positives ou négatives?
3- Autres suggestions, recommandations
4- Identification

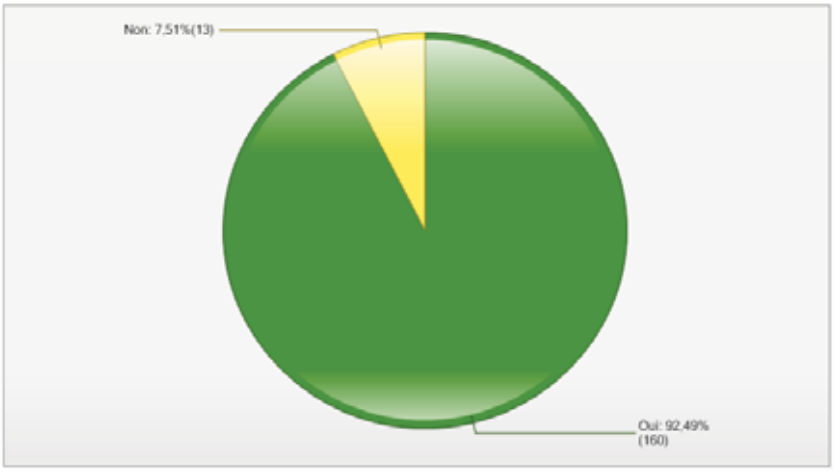
a. Société b. Secteur c. Catégorie d. Nom/prénom	e. Email f. Fonction g. Ouvert à contribution dans le cadre de la démarche en cours

3.3. Eléments de synthèse de l'enquête

Recherche

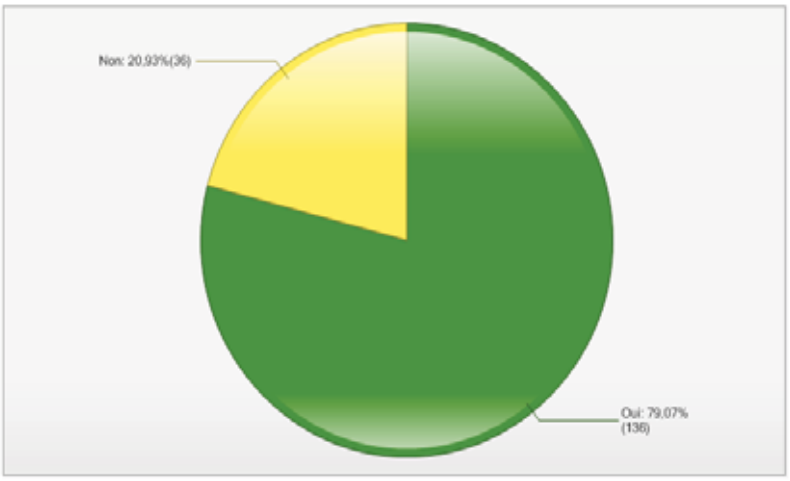
- Les chercheurs qui ont répondu à l'enquête, bien qu'ils ne représentent qu'une partie des chercheurs français travaillant sur des systèmes d'intelligence artificielle, estiment à plus de 90% que leurs résultats sont « suffisamment matures pour devenir une réalité industrielle à court terme ». Le momentum actuel est ainsi confirmé.

Pensez-vous que ces sous-domaines soient suffisamment matures pour devenir une réalité industrielle à court terme ?



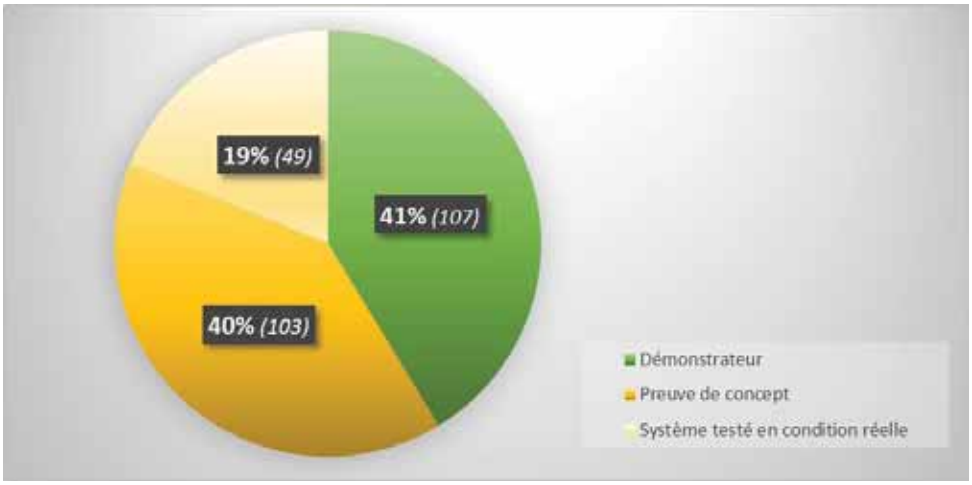
- Par ailleurs, ils estiment pouvoir transférer aux industriels le fruit de leurs travaux de recherche sous moins de 3 ans, pour près de 80% d'entre eux, reflétant une confiance suffisamment solide en l'applicabilité de ces résultats. Il est à noter que certains de ces « résultats de recherche » ne constituent pas, très légitimement, des produits transférables, dans le cas, par exemple, de travaux théoriques.

Disposez-vous de résultats que vous considérez susceptibles d'être industrialisés et transférés pour une utilisation opérationnelle à court terme (moins de trois ans) ?

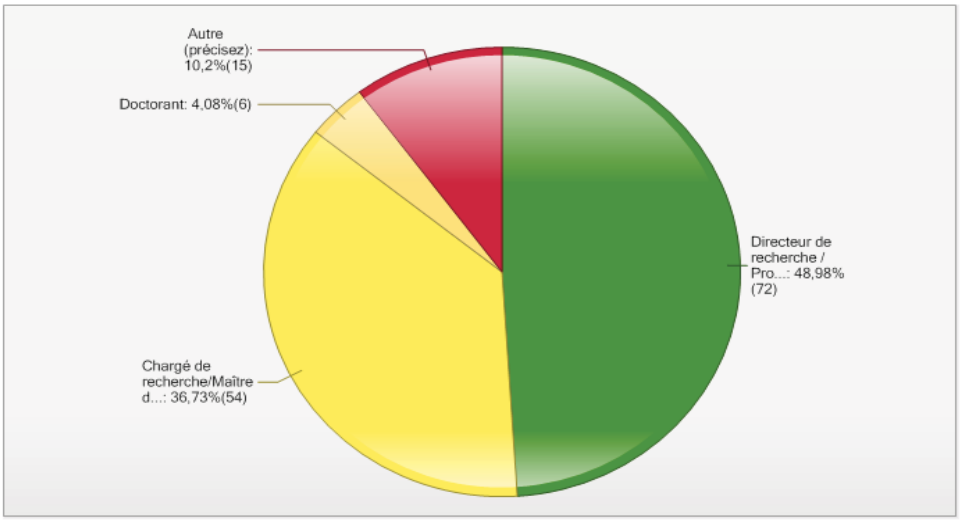


- Lorsque l'on interroge les chercheurs sur le degré de maturité de leurs résultats, ils ventilent ceux-ci selon 3 des états de pré-industrialisation : démonstrateurs, preuve de concept ou système testé en condition réelle. Plus de la moitié des résultats proclamés (59%) par les chercheurs sont donc d'ores et déjà prêts pour un usage industriel à court et moyen terme.

Qualifiez le degré de maturité technologique actuel de ce résultat



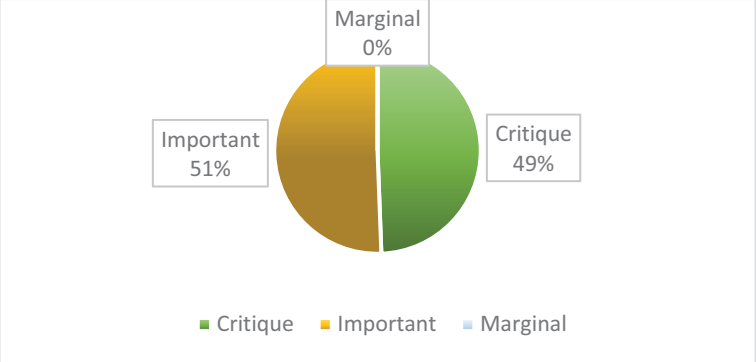
- Les répondants à l'enquête sont, sans surprise, majoritairement les directeurs de recherche ou les chercheurs eux-mêmes.



Entreprises

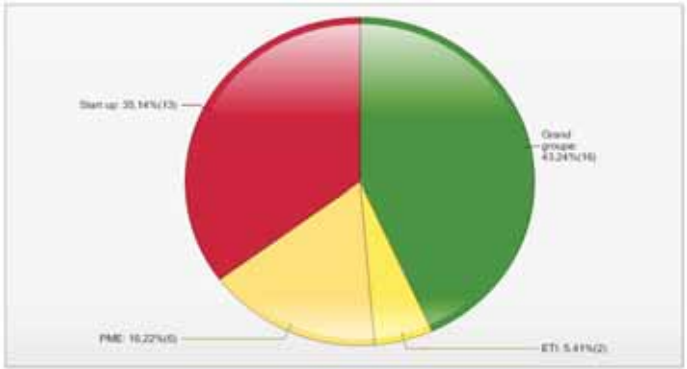
- Pour la cinquantaine d'entreprises qui ont répondu au questionnaire, les bénéfices attendus des technologies d'IA sont qualifiés comme importants voire même critiques.

Quels avantages apporterait une solution intégrant de l'IA ? Pourriez-vous qualifier sur une échelle de 1 à 3, l'importance du bénéfice envisagé (1 : marginal, 2 : important, 3 : critique),



- Les répondants sont issus de toutes les catégories d'entreprises

Origine des répondants



3.4. Détails de principales recommandations

Pour traiter les enjeux d'industrialisation et d'appropriation des résultats de la recherche en IA, le groupe de travail a formulé **sept recommandations principales** déclinées par ordre de priorité décroissant.
Ces dernières sont détaillées dans les pages suivantes.

8. Organiser des **challenges ouverts et des bourses aux technologies d'IA** répondant à des besoins des entreprises et de la société, **en complément des actions de fond** de montage de projets collaboratifs, de laboratoires communs et de structures agiles d'échange et de transfert pour soutenir la rencontre entre recherche et entreprises ; parmi les appels à projets, on inclura **un volet sur l'explication,** condition importante pour l'appropriation et le transfert des systèmes d'IA ;
9. Lancer un **grand projet public fédérateur en IA** comprenant notamment la réalisation d'un assistant conversationnel intelligent en soutien à la modernisation de l'action publique, mis en ligne sur les sites et les applications du service public ;
10. Soutenir la création de **plateformes communautaires, matérielles et logicielles**, dans quelques domaines prioritaires décrits ci-dessous ;
11. Confier au Laboratoire National de métrologie et d'Essais (LNE) la mission et les moyens de fédérer le **développement des méthodes de mesures, de tests et de certification pour les systèmes d'intelligence artificielle,** en partenariat avec les communautés de recherche correspondantes ;
12. **Dédier un fonds d'investissement en capital** (supérieur à 25M € par action) pour soutenir la **transformation des start-ups** dans le domaine en futurs champions ;
13. **Accélérer la mise à disposition de données publiques ouvertes pour les domaines stratégiques de la santé et de l'énergie**, afin de permettre le développement de solutions IA qui ont particulièrement besoin d'être alimentées en données réelles pour être efficaces et pour leur adoption par les principaux acteurs économiques ;
14. **Créer une Fondation de l'IA**, dont les modalités sont à définir, afin d'incarner et fédérer dans la durée l'écosystème français de l'IA autour des défis de la société et des entreprises.

Le groupe recommande également pour la question spécifique des **plateformes communautaires**, de se focaliser sur :

- La création d'une plateforme générique **d'intégration IA**, future composante de la plateforme européenne « AI-on-demand » prévue dans le prochain appel à propositions H2020 ;
- Une plateforme **« Données, logiciels et ressources de calcul pour l'apprentissage automatique** » ;
- Une plateforme **« véhicule autonome »** réunissant les briques matérielles (y compris piste d'essais, sites, véhicules) et logicielles (algorithmes et simulateurs) permettant de développer les technologies de perception et de prise de décision pour les véhicules autonomes de demain ;
- Une plateforme de **ressources pour le traitement automatique de la langue naturelle, l'interaction humain-machine** et les agents conversationnels.

1. Transfert recherche-entreprises

Recommandation n° 1
Organiser des **challenges ouverts et des bourses aux technologies d'IA** répondant à des besoins des entreprises et de la société, **en complément des actions de fond** de montage de projets collaboratifs, de laboratoires communs et de structures agiles d'échange et de transfert.
Résumé
Organiser des challenges ouverts et des bourses aux technologies d'IA permettant à une entreprise, en particulier les PME/ETI, voire à l'administration, de lancer un challenge auprès de start-ups ou solliciter le soutien d'organismes de recherche. L'objectif serait de répondre à une difficulté technologique ou évaluer l'intérêt potentiel (y compris d'usage) d'une technologie avant son implémentation. **Appels à projets « multipartenaires »** impliquant obligatoirement a minima un organisme de recherche et un industriel. Sélection par l'excellence et l'impact économique de projets de recherche industrielle utilisant des technologies existantes dans les laboratoires de recherche, pour les faire passer à la dimension industrielle, soit un passage de TRL 3 à 6 (i.e. *Technology Readiness Level,* niveau de maturité technologique*)*. **Appels à projets de type « laboratoire commun »** entre académique et entreprise dont l'objectif serait là aussi l'industrialisation de résultats de recherche ou de technologies. **Appels à projets sur l'explication ;** parmi les appels à projets, on inclura un volet sur l'explication des décisions**,** condition importante pour l'appropriation et le transfert des systèmes d'IA.
Evaluation d'impact et de coût
400M€ sur 10 ans, y compris la participation des entreprises à hauteur de 30 à 50% selon le niveau de TRL et la nature de l'entreprise.
Commentaires
Une condition nécessaire serait la démonstration rapide d'une plus-value mesurable pour un des secteurs utilisateurs, et la participation effective d'entreprises du secteur y compris au niveau du pilotage des projets.

2. Lancer un grand projet public fédérateur en IA « MarIAnne »

Recommandation n° 2
Grand projet public fédérateur «MarIAnne»
Résumé
Ce grand projet vise à développer notamment un assistant conversationnel intelligent en soutien à la modernisation de l'action publique, mis en ligne sur les sites de service public, utilisant les ressources en ligne et les nombreuses API réalisées par la Direction interministérielle du numérique et du système d'information et de communication de l'État (DINSIC).
Evaluation d'impact et de coût
50M€ y compris la contribution d'industriels intéressés par les marchés publics induits.
Commentaires
Cette initiative permettrait à nos services publics de ne pas dépendre des outils fournis par des acteurs étrangers pour l'exploitation de nos données publiques et personnelles gérées par les administrations, d'éviter la perte d'une part croissante de la valeur ajoutée à leurs profits et d'offrir une solution d'intégration des briques technologiques développées en France.

3. Plateforme Données, logiciels, ressources de calcul pour l'apprentissage automatique

Recommandation n° 3
Plateformes Données, logiciels, ressources de calcul pour l'apprentissage automatique
Résumé
Appels à projets pour des plateformes logicielles et/ou matérielles visant notamment **(1) le** parangonnage et à l'expérimentation voire l'organisation de challenges ou l'intégration et la démonstration des innovations. Ces plateformes auront principalement trois fonctions : o (i) fournir des données réalistes et représentatives des problèmes et usages réels, en quantité et en qualité, ainsi que des capacités d'exploitation et d'expérimentation – par exemple des sites de type « living labs » – permettant de tester des nouveaux services, des robots domestiques, des véhicules autonomes, etc. ; o (ii) fournir des ressources logicielles et matérielles permettant notamment le passage à l'échelle des méthodes et algorithmes sur les données mises à disposition ; o (iii) être équipées de capteurs, d'actionneurs, et d'analyseurs permettant de mesurer les résultats des expérimentations et ainsi de parangonner les technologies. **(2)** au développement de chaînes d'outils et de méthodes pour la conception, l'intégration ou la validation de briques logicielles et de systèmes complets
Evaluation d'impact et de coût
100M€
Commentaires
Accès aux entreprises privées selon des modalités à définir.

4. Développer une expertise nationale en matière de méthodes de mesures, tests et de certification

Recommandation n° 4
Développer une expertise nationale en matière de méthodes de mesures, tests et de certification
Résumé
Confier au Laboratoire National de métrologie et d'Essais (LNE) la mission et les moyens de fédérer le **développement des méthodes de mesures, de tests et de certification pour les systèmes d'intelligence artificielle,** en partenariat avec les communautés de recherche correspondantes.
Evaluation d'impact et de coût
A définir
Commentaires
Un enjeu fondamental pour le développement de l'IA en France est de mettre en place l'ensemble de la chaine d'outils (protocoles, méthodes, simulateurs, environnements de conception et d'intégration, etc.) permettant d'expliquer, de garantir et de certifier les technologies utilisant de l'IA. Il est proposé que cette action soit mise sous la responsabilité d'animation du Laboratoire National de métrologie et d'Essais (LNE) dont le domaine de l'intelligence artificielle est un des axes stratégiques de développement. Objectifs : - positionner la France en acteur central pour la définition et le contrôle des normes et standards du déploiement de l'IA ; - favoriser l'appropriation des solutions IA par les différents acteurs et la société civile ; - rendre les acteurs français de l'IA exemplaires en la matière et donc en faire des partenaires internationaux incontournables (à haute valeur ajoutée y compris pour les géants de l'IA).

5. Dédier un fonds d'investissement pour des start-ups en IA

Recommandation n° 5
Dédier un fonds d'investissement en capital pour le développement des start-ups spécialisées en IA
Résumé
D'un montant supérieur à 25M€ par action, ce fonds est destiné à soutenir la **transformation des start-ups** dans le domaine en futurs champions.
Evaluation d'impact et de coût
300M€ sur 10 ans
Commentaires
3e tour de table

6. Accélérer la mise à disposition de données publiques ouvertes pour les domaines stratégiques de la santé et de l'énergie

Recommandation n° 6
Accélérer la mise à disposition de données publiques ouvertes pour les domaines stratégiques de la santé et de l'énergie
Résumé
Initiatives spécifiques aux secteurs santé et énergie ; • Santé : soutenir la **constitution de bases de données médicales et de santé**, en étant attentif aux questions réglementaires et de vie privée ; mettre en place un fonds d'investissement dédié aux technologies d'IA pour la santé ; • Energie : développement d'un **avatar temps-réel du réseau** s'appuyant sur **une matrice de visibilité** de l'ensemble des données, afin de soutenir en particulier la mise en place de deux grands défis : le développement d'une **solution d'équilibrage**, et la conception de **places de marché** pour l'énergie.
Evaluation d'impact et de coût
A définir
Commentaires

7. Créer la fondation de l'IA

Recommandation n° 7
Créer la Fondation de l'IA
Résumé
Dans des modalités restant à définir, la Fondation vise à incarner et fédérer dans la durée l'écosystème français de l'IA autour des défis de la société et des entreprises. Elle sera un lieu privilégié d'échanges et de diffusion sur les avancées, potentialités, opportunités et risques liés à l'intelligence artificielle. Elle permettra aux chercheurs, au grand public et aux industriels de mieux appréhender les grandes questions posées par l'IA.
Evaluation d'impact et de coût
10M€ dont co-investissement industriel
Commentaires
Un dispositif focalisé uniquement sur l'IA est nécessaire pour accompagner la création et soutenir l'essor initial de la filière.

8. Créer une plateforme d'intégration et de démonstration des innovations #FranceIA

Recommandation n° 8
Créer une plateforme d'intégration et de démonstration des innovations #FranceIA
Résumé
La création d'une plateforme générique **d'intégration IA** constitue la future composante de la plate-forme européenne « AI-on-demand » prévue dans le prochain appel à propositions H2020 La plateforme « vitrine nationale » #FranceIA viserait à faciliter, sous une ombrelle commune et via un accès pour tous (industriels, chercheurs, entreprises, voire grand public) pour l'expérimentation, l'intégration et la démonstration en grande dimension des innovations et le parangonnage des technologies. Les plates-formes auront principalement deux fonctions : (i) fournir des données réalistes et représentative des problèmes et usages réels, en quantité et en qualité, ainsi que des capacités d'exploitation et d'expérimentation – par exemple des sites de type « living labs » – permettant de tester des nouveaux services, des robots domestiques, ou encore des véhicules autonomes; (ii) être équipées de capteurs, voire d'actionneurs, et d'analyseurs permettant de mesurer les résultats des expérimentations et ainsi de parangonner les technologies.
Evaluation d'impact et de coût
50M€
Commentaires
La question centrale à laquelle doit répondre cette recommandation de soutien de l'initiative nationale #FranceIA est la suivante : comment empêcher le siphonage des meilleurs par les GAFAM ? ou comment favoriser l'attractivité des entreprises françaises pour les jeunes chercheurs et ingénieurs : marque, rémunération, image. Il s'agit également de mettre en place des dispositifs à l'échelle nationale pour fédérer les communautés et préparer un possible « flagship » européen (dans l'hypothèse d'une plateforme « AI on demand » portée par la Commission européenne). Au travers de cet « espace » partagé entre tous les acteurs français de l'IA, il s'agit de stimuler la recherche nationale, favoriser la maturation des technologies et fédérer les acteurs de la filière au sens large (soit une approche intégrative incluant également l'implémentation dans les systèmes matériels des algorithmes d'intelligence artificielle) au sein d'écosystèmes entre recherche académique, technologique et entreprises. Cette plateforme et cette communauté doivent pouvoir bénéficier d'une équipe pérenne chargée de l'animation de l'écosystème et de l'identification des opportunités côté recherche et côté entreprises. Cette équipe sera notamment chargée de gérer le programme #FranceIA, le fonctionnement de sa gouvernance multi partenariale et assurer sa communication. Elle pourra s'appuyer sur les acteurs d'ores et déjà en place (type IRT, SATT) mais dédiée au programme #FranceIA.

9. Plateforme matérielle et logicielle « Véhicule autonome »

Recommandation n° 9
Une plateforme **« véhicule autonome »** réunissant les briques matérielles (y compris pistes d'essai, sites, véhicules) et logicielles (algorithmes et simulateurs) permettant de développer les technologies de perception et de prise de décision pour les véhicules autonomes de demain.
Résumé
Sur le plan matériel, cette plateforme comprendrait l'investissement dans des pistes d'essais et plateformes véhicules ; sur le plan logiciel, les simulateurs et boîte à outils.
Evaluation d'impact et de coût
A chiffrer par le groupe Véhicule autonome
Commentaires
Une quantité colossale de codes et de données en open source sponsorisés par les grands groupes (Udacity, Google, OpenAI) est disponible. Il s'agirait également d'y participer en accompagnant les initiatives des constructeurs français pour les positionner au mieux sur le secteur. En effet, le soutien du secteur nécessite un co-investissement public-privé conséquent.

10. Plateforme de ressources pour le traitement automatique de la langue naturelle, l'interaction homme-machine et les agents conversationnels

Recommandation n° 10
Une plateforme de **ressources pour le traitement automatique de la langue naturelle, l'interaction humain-machine** et les **agents conversationnels**
Résumé
Ouverte à tous secteurs d'activité, la plateforme de ressources mettrait à disposition des acteurs des textes annotés, traduits, étiquetés, des corpus de dialogues humain-machine, en langages naturels et multimodaux, etc. Elle permettrait également aux entreprises utilisatrices de réaliser des tests d'agents conversationnels.
Evaluation d'impact et de coût
50M€
Commentaires
En particulier, traiter parfaitement le français est une nécessité économique et culturelle ; on ne peut se reposer sur les acteurs anglophones.

11. Plateforme d'intégration et simulation

Recommandation n° 11
Une plateforme de **ressources pour l'apprentissage, et notamment le *deep learning***
Résumé
Plateforme matérielle et logicielle, par exemple autour des « architectures neuromorphiques ». Elle aurait pour objectif de développer toute la chaine d'outils nécessaires, allant de la conception d'applications à l'implémentation adaptés aux matériels, tenant compte des contraintes de consommation, de puissance du matériel, etc.
Evaluation d'impact et de coût
10M€ + 5M€ fonds d'amorçage Start-ups
Commentaires

12. Plateforme d'intégration multi-agents et Internet des Objets

Recommandation n° 12
Plateforme d'intégration multi-agents et Internet des Objets
Résumé
D'un point de vue multi-agent (plus généralement de la coopération, coordination d'agents encapsulant des technologies développées dans divers domaines de l'intelligence artificielle), il y a besoin d'une plateforme pour le parangonnage à grande échelle dans un cadre ouvert, hétérogène et décentralisé d'algorithmes de coordination et de régulation d'agents autonomes : agents développés par des opérateurs différents, déployés dans un environnement dynamique et distribué en interaction avec besoin de satisfaire des buts communs ou non. Une extension à l'Internet des Objets est pertinente dans la mesure où l'approche multi-agent permet de traiter les questions de communication, de coordination, de prise de décision collective d'objets connectés « simples » mais en très grand nombre, ou d'objets « lourds » dotés de capacités de raisonnement.
Evaluation d'impact et de coût
20M€
Commentaires
Il sera possible de décliner cette plateforme sur différents domaines applicatifs (énergie, transport, véhicules autonomes, pilotage d'ateliers de production, etc.), voire de les combiner (p.ex. pour la ville intelligente).

GROUPE DE TRAVAIL 2.2
DÉVELOPPEMENT DE L'ÉCOSYSTÈME

Groupe Développement sur le territoire national de l'écosystème français des fournisseurs français de technologies de l'IA et de leur utilisation

Périmètre d'étude du groupe de travail

Le groupe de travail 2.2 est composée de personnalités d'origines variées permettant de couvrir tant les aspects technologiques que sectoriels. D'un commun accord, les participants ont souhaité orienter les travaux autour de deux aspects complémentaires : d'une part les conditions techniques pour le développement des entreprises en lien avec l'Intelligence Artificielle, et d'autre part les questions relatives au déploiement des pratiques d'usage dans les différents secteurs de l'Economie..

Ne pouvant faire une étude intégrale de l'ensemble des problématiques rencontrée par les entreprises (start-ups en création, PMEs et GE) dans tous les domaines, nous avons choisi d'axer notre réflexion sur quelques thématiques choisies afin d'illustrer les difficultés et les bonnes pratiques rencontrées dans quelques domaines de l'IA. Par ailleurs, nous avons également effectué ces travaux sur une Région, choisie en raison de sa taille, de sa diversité et donc du fort impact des technologies de l'IA sur le développement de ses entreprises.

A travers ces exemples, nous avons identifié quelques enjeux et points de blocages qui nous ont paru les plus couramment rencontrés dans le développement des entreprises de l'IA et de celles qui l'utilisent.

Etat des lieux

Le déploiement de l'Intelligence Artificielle dans les entreprises est directement corrélé à la disponibilité des données, aux infrastructures numériques permettant de les traiter et à la confiance dans la pérennité de ces outils.

Le succès du déploiement de l'Intelligence Artificielle repose sur 3 piliers techniques ou technologiques:

- l'algorithmique (sujet déjà traité par les groupes de travail recherche) ;
- les technologies des infrastructures *scalables* (computer, storage, networks) pour assurer le traitement, le stockage et les échanges des données massives ;
- l'accès à des bases de données massives brutes (ou prétraitées) et annotées de façon à faciliter l'apprentissage et l'utilisation par des usagers non techniques.

Infrastructure matérielle de Calcul Haute Performance pour l'IA.

La complexité croissante des algorithmes d'apprentissage machine, notamment de Deep Learning, couplée à l'amélioration des performances **avec le volume de données utilisées,** rendent indispensable la disponibilité d'infrastructures de calcul puissantes, souvent distribuées**. L'usage des technologies de « cloud computing » est souvent présenté comme la solution, mais il pose deux défis majeurs** :

- Les « clouds » actuels mettent essentiellement en œuvre des serveurs de calculs d'architecture classique dite CPU, alors que les processeurs GPGPU (General Purpose Graphical Processing Unit) s'avèrent offrir une puissance de calcul très nettement supérieure tant pour la phase d'apprentissage que pour l'exécution des modèles appris, pour des couts de machines comparables, et même parfois une efficacité énergétique accrue. De grandes capacités de mémoire vive (RAM) sont aussi nécessaires. Ainsi, les grands acteurs de « Clouds » américains (comme Amazon Web Services) ont commencé à intégrer ces processeurs (acquis auprès de NVIDIA, acteur américain leader du marché des GPU).
- **le besoin d'un « cloud computing » souverain**, au moins au niveau européen, localisé sur le territoire européen et opéré par des acteurs européens de confiance, condition sine qua non pour garantir la sécurité des données qui y sont déposées et traitées.

Accès et utilisation des données :

Les données constituent aujourd'hui l'un des carburants majeurs de l'IA. Celles-ci sont notamment utilisées par des algorithmes d'apprentissage machine qui produisent des modèles mathématiques qui sont ensuite utilisés pour des prises de décision. Ces algorithmes peuvent apprendre sur des données brutes, en **mode supervisé ou non** et produisent des modèles prédictifs à partir de données annotées par des valeurs cible. De plus, certaines approches d'apprentissage très performantes ont besoin d'un nombre de données importantes.

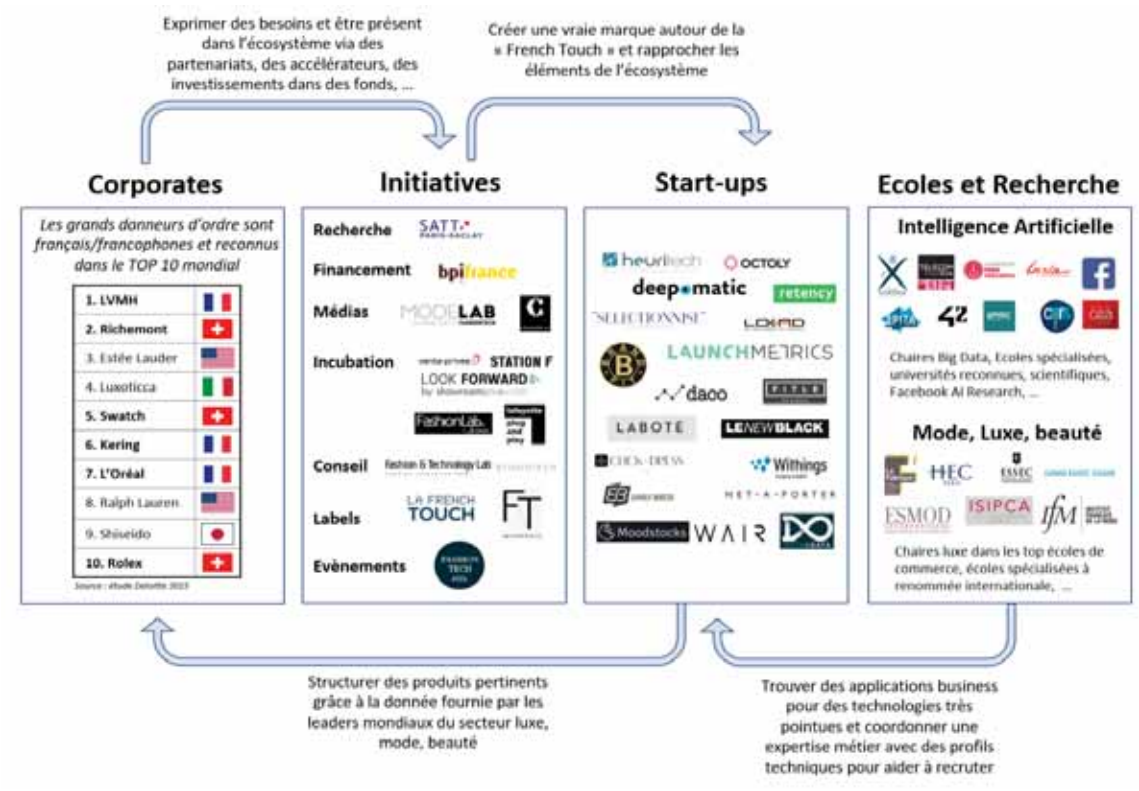
Un 1er écueil majeur, pour toutes les entreprises travaillant sur l'IA, concerne **l'accès et l'utilisation effectifs des données** pourtant produites en masse par différentes organisations (entreprises, administrations, ..). **S'agissant des données à caractère personnel, il existe d'une part un risque juridique et financier fort** lié au caractère uniquement contraignant et répressif du règlement européen 2016/679 (GDPR), qui, outre des mesures de sécurité et organisationnelles légitimes, impose des restrictions encore plus fortes sur l'utilisation de telles données, notamment en termes de finalité et de durée, ce qui limite très fortement les approches exploratoires et donc la recherche et l'innovation. On constate de plus un risque d'image important pour les entreprises concernées lié à une opinion publique de plus en plus sensible sur le sujet des données et de l'IA en général, et nettement moins tolérante à l'égard des acteurs français que des acteurs US pourtant moins respectueux en la matière. **Concernant des données techniques** produites lors des opérations d'infrastructures, de véhicules, etc., **la difficulté d'accès est plus liée à la confidentialité ou à la sensibilité « business ».** La possession de telles données peut en effet permettre de « désintermédier » et de donner les clés business à de nouveaux acteurs, en particulier ceux qui maîtrisent le stockage, les flux et l'accès aux données et aux utilisateurs (GAFAM).

Un 2nd écueil concerne la production d'annotations de données brutes, qui peut s'avérer extrêmement couteuse, voire très difficile à réaliser, car requérant très souvent une expertise métier pointue, difficulté amplifiée par la volumétrie importante nécessaire. De façon critique, quel que soit le modèle des entreprises qui travaillent sur ces données, la valeur ajoutée, pour l'entreprise consiste en la collecte, la mise en forme et l'analyse (par le l'IA) de ces données.

La connaissance du domaine d'application et l'accès à des utilisateurs potentiels permettant d'expérimenter et d'évaluer de nouveaux usages conditionnent la capacité de déploiement de l'IA.

Nous traitons ici certains domaines en mettant en évidence des différentiateurs, mais surtout les points de blocage à résoudre pour pouvoir assurer le succès de l'écosystème au niveau national.

Luxe - Mode :



Structure de l'écosystème dans le secteur de la mode

La France a la chance de regrouper les **deux types d'acteurs essentiels** à la création d'une vraie filière autour de l'Intelligence Artificielle dans le secteur Luxe-Beauté-Mode : des leaders mondiaux qui sont à la recherche d'innovation pour rester à la pointe (LVMH, l'Oréal,…), des start-ups en croissance (Heuritech, deepomatic, LaunchMetrics…) et des centres de recherche et écoles tels Epitech présent dans notre groupe, qui forment des talents prêts à s'investir sur ce marché. Un enjeu sera de favoriser l'émergence de startups dans ce secteur en formant des ingénieurs "FashionTech" ayant la double compétence *datascience* / *machine learning* ainsi qu'une connaissance du monde Luxe-Beauté-Mode. Il s'agit également de s'appuyer sur le label « French Touch » pour le matérialiser plus autour de la « FashionTech » et de l'AI afin de ²mettre en valeur et de vendre cet écosystème en dehors de la France.

Assurance :

L'Assurance est certainement l'un des secteurs les plus concernés par l'opportunité de l'intelligence artificielle et du traitement de données. D'abord, son principe central – la mutualisation du risque – repose en effet directement sur les théorèmes fondamentaux de statistiques. Les produits d'assurances sont ainsi tous construits grâce à l'analyse statistique des données de sinistralité. Ensuite, l'Assurance touche à de nombreux secteurs de l'économie, si bien que l'impact des nouvelles technologies sur ce secteur en est démultiplié. Le développement des objets connectés (et **des intelligences artificielles associées**) impacte par exemple, l'assurance automobile avec les véhicules connectés, l'assurance

habitation avec la domotique ou encore l'assurance santé avec le « quantified self ». Enfin, devant gérer la relation avec un grand nombre de clients et traiter un volume important de sinistres, elle présente naturellement une formidable opportunité d'application de l'intelligence artificielle, que ce soit pour optimiser et automatiser ses processus ou pour s'armer de capacités prédictives plus poussées sur des sujets comme l'attrition ou la détection de la fraude.

Les assureurs français ont très tôt pris conscience de leur position et pris des mesures pour être acteurs de leur transformation. Deux approches non exclusives sont principalement adoptées : le développement en interne de nouveaux produits par des équipes spécialisées et le travail avec des start-ups. Les projets se rangent naturellement dans deux catégories :

- La première est l'utilisation des nouvelles données disponibles pour créer de nouveaux produits d'assurances. Un exemple intéressant est celui de la tarification basée sur des données de télématique automobile, pour lequel le groupe AXA a organisé une compétition de *machine learning* destinée aux meilleures solutions.
- La deuxième catégorie de projet concerne l'amélioration d'un maillon existant de la chaîne de valeurs. Dans le cas de la détection de la fraude par exemple, de nombreux assureurs ont choisi la solution de la start-up française Shift Technology spécialisée sur le sujet.

L'utilisation de l'intelligence artificielle en assurance se fait donc sans blocage particulier mais elle prend du temps. Pour être effective, celle-ci doit en effet s'inscrire dans une transformation globale des processus et modes de gestion, ce qui nécessite un travail important d'accompagnement du changement. Enfin, l'application des technologies aux données extrêmement spécifiques que sont les données d'assurance nécessite une véritable expertise métier, qui ne s'acquiert qu'avec des années d'expérience.

Finance :

Dans le secteur de la Finance, un écosystème naît autour de l'IA mais reste à fortement renforcer. Notre réflexion s'est portée sur un cas concret avec la mise en place d'un « chatbot » Do You Dream Up dans une banque française. Ce « chatbot » doit servir de support aux employés sur des questions récurrentes relatives à leurs postes de travail.

Les résultats attendus sont les suivants : gain de productivité, **immédiateté** de résolution de problème, autonomie et montée en compétence des agents de support, dès lors libérés de tâches répétitives.
Les constats post-déploiement sont les suivants : facilité d'allocation de ressources pour la mise en œuvre, via du sponsoring de hauts dirigeants et grâce à la volonté d'innover propre au secteur. Cependant, de nombreuses difficultés sont observées en raison de l'agilité peu pratiquée dans ces structures, parfois très contraignantes (niveau sécurité et organisationnel). Par ailleurs, l'importance des fausses croyances autour de l'IA ne permet pas de déployer une stratégie globale mais force à agir au "fil de l'eau".

Le panorama des startups françaises qui sont « data-driven » avec des briques d'intelligence artificielle, ne fait pas apparaître de secteur leader.

En faisant un panorama des startups françaises (réalisé par Serena Capital premier fonds européen dédié au Big Data et à l'IA dans les initiatives privées autour de l'IA) qui sont « data-driven »[1] avec des briques d'intelligence artificielle, il apparaît qu'il n'y a pas vraiment de secteur leader.

En effet, le marché adressé par l'intelligence artificielle est très fragmenté, même si environ un tiers regroupe trois secteurs clefs :

- Le logiciel : surtout des logiciels B2B qui vendent à de grandes entreprises, qui ont les data sets suffisants pour pouvoir faire tourner des modèles et qui ont besoin de valoriser cette donnée dans tous leurs processus
- La fintech/regtech/insurtech : pour fournir des services à plus forte valeur ajoutée, les startups se servent de l'intelligence artificielle pour le support client (les chatbots), l'offre produit (plus personnalisée), prévenir la fraude, etc.
- Le marketing : la tendance est à la forte personnalisation des produits, et une forte attention portée à ce qui est dit sur un produit sur les réseaux sociaux ou alors la détection de tendances.

Pour les 2/3 restants, de nombreux verticaux sont adressés : la santé, l'éducation, les ressources humaines, le transport, la sécurité, les smart cities/smart homes, l'agriculture, etc.

Le développement d'une thématique régionale sur l'Intelligence Artificielle est particulièrement pertinent, en ce qu'elle touche de multiples secteurs et entreprises : l'exemple régional de l'Occitanie.

La région Occitanie, via ses différentes structures d'accompagnement (DIRECCTE, IRT, pôles de compétitivité) soutient une démarche novatrice qui croise ses filières d'excellence avec les nouvelles opportunités qui s'ouvrent comme l'intelligence artificielle, le traitement des données massives et l'internet des objets (détails en ANNEXE 2).
La région comprend des acteurs industriels de premier plan (Continental, Airbus, Sigfox), tous grands pourvoyeurs de données massives et qui ont engagé une démarche forte en IA (projet ambitieux de collecte massive de données à bord des véhicules chez Continental, maintenance prédictive chez Airbus avec une « académie des analytics », pour former les talents internes. Par ailleurs, la présence du Centre national de recherches météorologiques, du principal centre de recherche INRA dans le domaine de l'analyse des données agricoles et de l'intelligence artificielle permettent également d'ouvrir l'IA à d'autres secteurs comme la météo ou l'agriculture et de renforcer cet écosystème régional.
Cette région voit ainsi dans la transformation numérique et la valorisation de ces données des opportunités stratégiques qui méritent d'être traitées de concert, afin créer un effet d'entraînement. Le développement des activités en lien avec l'IA, notamment avec de la formation interne, doit permettre une forte structuration de l'écosystème régional.

[1] Source : étude 2016 deal flow data de Serena Capital (https://blog.serenacapital.com/data-driven-startups-deal-flow-is-rising-e1ce3d1d1908#.s86n92tze)

Enjeux et recommandations

Recommandation n°0 : Etendre à l'Intelligence Artificielle le périmètre stratégique des données. (Actions PIA3 : Concours d'Innovation ciblé)

La France a développé une stratégie autour de la donnée comprenant une des 9 solutions de la stratégie Nouvelle France Industrielle ; des outils comme le Concours Mondial de l'Innovation (Ambition 7), les Challenges Big Data et Numérique, les Grands Défis du Numérique, les différents appels PIAVE et FSN (Générique et Boosters notamment) ; des plateformes comme la plateforme Teralab ; des plans d'action Européens (PPP Big Data) ou binationaux (Feuille de route France-UK autour de la donnée, action Franco-Allemande) et des formations dédiées.

Cependant, que ce soit pour des données en stock ou en flux, et quelle que soit leur modalité (textuelle, vocale, image, vidéo, vision, etc) ou nature (personnelle, de santé, fournie par des capteurs ou non,…), les traitements algorithmiques sont liés aux données. Nous recommandons donc d'étendre explicitement à l'intelligence artificielle la stratégie Economie de la Donnée, en accroissant les moyens des différents instruments sans en inventer des nouveaux.

1. Les premiers enjeux sont techniques et consistent à mettre à disposition des acteurs industriels les infrastructures et les données leur permettant de mettre en place des actions durables

Face aux géants du Numérique, l'accès aux données massives et de multiples origines est un impératif. Pour **les données personnelles**, le risque juridique et financier soulève un fort enjeu à compléter l'encadrement du traitement des données à caractère personnel par un volet plus positif par exemple en identifiant des catégories de traitements légitimes au regard de finalités scientifiques, technologiques ou économiques et à éduquer l'opinion publique par rapport à ces enjeux. A titre d'exemple, citons les données médicales personnelles (pathologies, examens, imagerie) qui relèvent directement de cette catégorie de données et dont l'enjeu en termes de progrès médical est considérable. **Sur les données techniques**, afin d'obtenir l'accès à de telles données pour innover sur leurs usages, il est important de se doter de conditions qui en garantissent la confidentialité. Ceci requiert différentes initiatives à favoriser telles les plateformes d'expérimentation proches – voire sous contrôle – des producteurs ou des propriétaires de ces données.

Enfin, sur la production d'annotation sur les données brutes, l'**enjeu est d'arriver à aider sur le plan financier, juridique et organisationnel la production et l'utilisation par les acteurs industriels et de la recherche de tels corpus de données annotées.** Ces annotations constituent bien souvent un des cœurs de métier des entreprises de l'IA. Une attention toute particulière devrait être accordée aux données langagières écrites et orales de par l'importance considérable des applications tant pour l'interaction homme- machine que pour l'exploitation des connaissances contenues dans les textes. La langue française particulièrement en retard, constitue ainsi une formidable opportunité de différentiation pour des acteurs français.

Sur le plan technique de la gestion de ces données, la structuration d'un cloud européen, et son équipement en outils de développement dédiés au Machine Learning offrant une grande flexibilité aux développeurs tout en assurant la portabilité des applications est également une des priorités.

La France dispose d'atouts évidents et reconnus tant sur les outils scientifiques et technologiques que sur les composants les systèmes d'intelligence artificielle que sur l'existence de PMEs et Startups susceptibles de déployer ces outils dans l'économie.

Pour rendre l'Intelligence Artificielle efficace il faut disposer d'une masse importante de données si possible annotées. Certains secteurs industriels et grandes entreprises disposent de ces données, mais celles-ci ne sont pas en général publics et concernent parfois des domaines privés, stratégiques …

Recommandation n°1 : Faciliter l'accès aux jeux de données et aux moyens de traitement pour le développement d'Intelligences Artificielles par filière et dans les régions. (Action PIA3 : AAP collaboratifs ciblés (PSPC), plateformes mutualisées sur une filière)

2. L'émergence de start-ups en IA dans les dernières années est notable, cependant certains domaines d'application demeurent à la marge.

La cartographie réalisée au cours de cette étude met cependant d'ores et déjà en avant des manques sectoriel dans l'émergence des start-ups IA : les secteurs plus caractéristiques de l'IA (robotique, big data) sont bien couverts, en revanche le déploiement à d'autres domaines se fait plus lent (ex : moins de 10 start-ups au total en sécurité, maison connectée, finance, assurance, administration, ...). Sur les 3500 petites entreprises aidées par Bpifrance chaque année, seules ~50 se réclament de l'IA en 2015, un chiffre relativement faible au vu de l'intégralité des secteurs couverts par l'IA. Ainsi, l'offre de technologie n'est pas toujours au rendez-vous, pas assez mature et éventuellement difficile à intégrer. Il y a donc un fort enjeu sur l'émergence de start-up dans des domaines a priori moins couvert par l'IA en France (sécurité, détection de comportements anormaux par exemple).

Si l'intelligence artificielle s'applique bien à tous les verticaux, l'enjeu réside dans l'expertise métier que peuvent avoir les startups sous-jacentes. Il est donc important de pouvoir trouver des *data-scientists* orientés métiers pour pouvoir vraiment adapter leurs algorithmes à des problématiques opérationnelles. Un autre enjeu concerne la maturité des secteurs concernés : par exemple, l'agriculture en général souffre d'un déficit de digitalisation par rapport au marketing, donc l'intelligence artificielle mettra sûrement plus de temps à se propager.

Recommandation n°2 : Diffuser les méthodes de l'Intelligence Artificielle grâce à une verticalisation sur des thématiques métiers en créant des Hubs IA (« mini cluster » réunissant la recherche, l'enseignement, l'incubation, l'accompagnement et le financement). **(Action PIA3 : AAP Nouveaux écosystèmes d'innovation).**

Recommandation n°3 : Un programme pour détecter, soutenir et accompagner les talents entrepreneuriaux issus du monde étudiant et universitaire, susceptibles de développer un projet de startup IA.

3. Dans les domaines où les startups en IA commencent à émerger, l'enjeu majeur de déploiement consiste à démystifier l'IA, à faire la pédagogie de ses enjeux et à faire comprendre la proposition de valeur pour les entreprises.

En premier lieu, ressort un besoin transversal d'élaborer un référentiel de l'IA partagé par tout l'écosystème qui permette à tous de parler le même langage et de faire de la pédagogie : glossaire, exemples concrets sectoriels, type d'applications / services, aspects juridiques et PI, consultant ressources, labos de recherche, lieux d'expérimentations pour les entreprises, cartographie de l'offre industrielle offre produits services pour quel enjeu, lieux et « cursus » de formation professionnelle,

diplômes d'ingénieurs intégrant de la RA, type d'accompagnement. Ainsi, les besoins consistent surtout à **démythifier l'IA et à en faire de la pédagogie sur les enjeux**, valeur pour les entreprises, risques emploi, sécurité, PI, etc…

Recommandation n° 4 : Mise à disposition d'un « guide d'identification des tâches professionnelles automatisables ».

Par ailleurs, il subsiste de nombreuses réticences pour mener des expérimentations qui utilisent des techniques d'IA. L'IA en encore est à ses prémices, avec des objectifs et gains pas suffisamment quantifiables. Les expériences menées ne sont pas encore intégrées dans la plupart des cas aux process classiques de l'entreprise (Très/trop peu de demandes dans les aides à l'innovation couramment rencontrées) :

- Besoin de démontrer la proposition de valeur rendue possible grâce à l'IA en explicitant le modèle économique qui en résulte (toutes entreprises).
- Donnant plus de place aux expérimentations en partenariat (toutes entreprises).
- Montrant les usages de l'IA aux entreprises (lieu vitrine et test ou action spécifique de démos pour les entreprises), en particulier pour les **PME et l'Administration, qui semblent éloignés de la problématique.**

Recommandation n°5 : Etendre au domaine de l'Intelligence Artificielle le Diagnostic Big Data proposé par Bpifrance

Recommandation n°6 : Faire des services publics locaux des terrains d'expérimentations avant d'envisager des déploiements nationaux. (Action PIA3 : **AAP soutien à l'innovation collaborative**).

Recommandation n° 7 : Défiscaliser le rachat de startups en IA.

Dans les secteurs où les outils technologiques de l'IA existent, ont été testés sur des cas test depuis plusieurs années et ont atteint une maturité leur permettant d'être largement déployés sur les domaines applicatifs divers. L'enjeu actuel est donc de déployer le plus largement possible.

Enfin, l'IA étant une thématique multi-sectorielle par construction, le développement régional de compétences en IA et de pôles leaders sur le sujet apparaît également comme un enjeu majeur du secteur.

4. <u>Lorsque l'IA est identifiée comme un enjeu majeur (startup, PME, GE), la question des compétences et des ressources internes et au niveau régional devient majeure.</u>

Cette thématique se retrouve tant pour les startups (Garder les compétences d'excellence en France, être compétitif /GAFA, inciter les ingénieurs IA à intégrer les startups françaises) que pour les GE (Airbus Helicopters a choisi de développer ses propres modèles mathématiques, et dans cette optique a mis en place une « académie des analytics », pour former ses talents internes).

Recommandation n°8 : Inscrire explicitement la thématique « Intelligence Artificielle » comme priorité des outils d'accompagnement du PIA2 et surtout du PIA 3 en cours de définition, notamment afin de faire émerger la thématique dans les domaines où elle n'apparaît pas, tout en renforçant son déploiement dans les secteurs où elle se développe. (PIA2 et PIA3 : **AAP Concours Mondial d'Innovation, AAP Concours Numérique, AAP PIAVE, AAP Disrupt' Campus**).

Recommandation n°9 : Sur un territoire, fédérer l'ensemble des acteurs concernés par l'IA et le Big Data. Ces acteurs travaillent ensemble afin d'établir une feuille de route ayant pour objectif principal de relever le défi de l'irruption de l'IA et plus, largement, de l'économie des données. L'ambition est renforcer le tissu de la formation et de la recherche et de construire un ensemble de cas d'usage qui seront autant de voies nécessaires à la pérennisation de ces filières. (**PIA3 : Nouveaux territoires d'innovation**).

5. En tant que technologie de traitement des données, l'IA acquiert immédiatement un caractère mondial et suppose donc un rayonnement national et international des entreprises du secteur.

Dans des domaines où la France dispose de leaders mondiaux (Luxe, Santé, Assurance,…), le besoin de visibilité mondiale des entreprises de l'IA est impératif pour se développer (pour recruter, pour gagner des contrats,…). L'effort essentiel de soutien à l'IA doit être focalisé sur la visibilité à donner aux start-ups du domaine (notamment pour le recrutement international), l'attraction des clients.

Recommandation n°10 : Construire, en capitalisant sur l'existant, une initiative ambitieuse pour animer l'écosystème IA français et le marketer à l'international afin de positionner la « French Tech » comme un des écosystèmes de référence dans le monde en matière d'IA.

Recommandation n°11 : Faire de la France le plus grand accélérateur virtuel jamais créé. Pas de lieux physiques où incuber les projets, tout est fait en ligne. **(Actions PIA3 : AAP Nouveaux écosystèmes d'innovation).**

Annexes

Annexe 1 – Contributeurs au groupe de travail 2.2

1. Structures d'accompagnement
    a. Pôles de compétitivité

- Image et Réseaux Henri SANSON henri.sanson@orange.com
- Cap Digital Philippe Roy Philippe.roy@capdigital.fr et Françoise Colaitis francoise.colaitis@capdigital.com.
- Systematic Najah Naffah nnaffah@prologue.fr
- Aerospace Valley Agnes Paillard agnes.paillard@eads.net

    b. French Tech Hub en France

- Directeur French Tech David Monteau david.monteau@finances.gouv.fr
- French Tech Toulouse Philippe Costes philippe.coste@epitech.eu

    c. SATT

- SATT Lutech : Damien Bretegnier damien.bretegnier@sattlutech.com

    d. IRT

- SystemX Paris Directeur Général Eric Perrin-Pelletier eric.perrin-pelletier@irt-systemx.fr
-

2. Financeurs
    a. Bpifrance

- Groupe de travail piloté par Laure Reinhart Responsable des partenariats Laure.reinhart@bpifrance.fr
- Anne Darnige, responsable du domaine expertise Numérique anne.darnige@bpifrance.fr
- Jean-Christophe Gougeon, responsable sectoriel Numérique jc.gougeon@bpifrance.fr
- Youri Jedlinski, responsable du Hub, youri.jedlinski@ext.bpifrance.fr

    b. Fonds Equity

- Serena Capital : Amélie Faure afaure@serenacapital.com
- Serena Capital : Lea Verdillon lverdillon@serenacapital.com

3. Acteurs
    a. Experts – PQ

- François Bourdoncle – Co fondateur Exalead francois.bourdoncle@fbcie.com

    b. Start ups

- Shift technology : Eric Sibony eric.sibony@shift-technology.com
- Heuritech : Tony Pinville pinville@heuritech.com
- DoYouDreamUp : Souad ZARAT s.zarat@doyoudreamup.com
- Davi : Pascal Arbault pascal.arbault@davi.ai
- Algolia : Liam Boogar liam@algolia.com

4. Administration
    a. DGE Groupe de travail piloté par Matthieu Landon Matthieu.landon@finances.gouv.fr
    b. CGI Laurent Rojey laurent.rojey@pm.gouv.fr

Annexe 2 - Recommandations des membres du groupe de travail 2 .2

Recommandations n°0

Titre : Extension au périmètre de l'Intelligence Artificielle de la stratégie autour des données

Résumé : Depuis plusieurs années, la France a développé une stratégie autour de la donnée comprenant :

- Une des 9 solutions de la stratégie Nouvelle France Industrielle
- Des outils comme le Concours Mondial de l'Innovation (Ambition 7), les Challenges Big Data et Numérique, les Grands Défis du Numérique, les différents appels PIAVE (Générique et Boosters notamment)
- La plateforme Teralab
- Des plans d'action Européens (PPP Big Data) ou binationaux (Feuille de route France-UK autour de la donnée, action Franco-Allemande)
- Des formations dédiées

Aujourd'hui, que ce soit pour des données en stock ou en flux, et quelle que soit leur modalité (textuelle, vocale, image, vidéo, vision, etc) ou nature (personnelle, de santé, fournie par des capteurs ou non,…), les traitements algorithmiques sont liés aux données. Nous recommandons d'étendre explicitement à l'intelligence artificielle la stratégie Economie de la Donnée, en accroissant les moyens des différents instruments sans en inventer des nouveaux.

Coût : Cout de mise en œuvre faible (extension du périmètre d'outils existant)

A compléter par une augmentation potentielle des enveloppes de ces outils – principalement en subventions et avance récupérables

Faisabilité/Maturité : 5

Impact : 4 ou 5 à chacun des critères (outils existant)

Commentaire : le rapport de France Stratégie[2], démontre une trop grande multiplication des outils de financements. Cette proposition vise à contenir cette inflation en rapprochant le soutien de l'Intelligence Artificielle de celui de l'économie de la donnée.

Recommandations n°1

Titre : Plateformes Collaboratives de Partage : Faciliter l'accès aux jeux de données et aux moyens de traitement pour le développement d'Intelligences Artificielles par filière et dans les régions.
(Action PIA3 : AAP collaboratifs ciblés (PSPC), plateformes mutualisées sur une filière)

Résumé : L'accès à des données brutes et annotées de différentes natures ainsi qu'à des moyens de calcul et de stockage pour l'exécution d'algorithmes d'apprentissage, demeure un verrou pour le développement de l'intelligence artificielle. Une solution consiste en la mise en place de *Plateformes Collaboratives de Partage* basées sur un modèle de Clouds dédiés, disposant des briques essentielles au cycle complet de traitement des données : de la capture à la prise de décision, permettant une

[2] Quinze ans de politique d'innovation, France Stratégie, 2016

exploitation contrôlée proche de l'endroit où les données sont produites (Edge Computing Haute Performance), éliminer les problèmes de latence et encourager l'innovation par filière professionnelle et par région. Ces plateformes réuniraient par filière professionnelle, détenteurs de données et de problèmes, chercheurs du monde académique et industriels pour lancer des projets d'innovation sur des durées variables (6 à 36 mois), tout en traitant les problèmes légaux et sécuritaires.

Coût :

Le coût de ces plateformes est variable selon la filière choisie. Des plateformes ou projets existants tels que Teralab, filière Science, et projets SDIL - Smart Data Innovation Lab (mécanisme proposé par l'association BDVA) donneront une estimation des investissements nécessaires (voir le lien http://www.sdil.de/en/platform/ qui illustre un exemple de technologies mises en œuvre).

Dans tous les cas, des subventions de service public (sous forme d'Appels à Projets) sont à prévoir pour amorcer la création de 6 plateformes réparties dans les régions, associées à un financement de la part des acteurs (détenteurs de données et de problèmes, et des industriels fournisseurs d'éléments d'infrastructure IT en mode Cloud)

Faisabilité/Maturité : 4

Un travail conséquent a déjà été fait, autour de la Big Data pour lancer des plateformes à usage ouvert. Citons en particulier Teralab et les SDIL dans le cadre de BBVA… qui pourront être valorisés pour le lancement des premières actions.

Néanmoins, Il demeure plusieurs questions de recherche :

- Répartition des traitements vers les données, grâce aux technologies de virtualisation (ex : containers) et la standardisation d'APIs.
- Garantir que les données privées des organisations les mettant à disposition ne pourront être récupérées que par les acteurs souhaitant les exploiter pour mettre au point leur IA, tout en permettant leur découverte et leur évaluation humaine, impliquant une chaine sécurisée de bout en bout et éventuellement l'application des techniques d'anonymisation.
- Capacité à exploiter conjointement des jeux de données qui restent cloisonnés dans des silos étanches (ex : apprentissage ubiquitaire, combinaison de modèles de connaissances extraits séparément, …).

Priorité : 5

Sans la mise en place de telles plateformes au niveau national et européen, en garantissant une autonomie et une liberté dans la gestion des déploiements et l'exploitation des résultats, nos institutions resteront dépendantes des fournisseurs de Clouds américains, et bientôt chinois. Ceci ne veut pas dire qu'il faudra s'enfermer dans des solutions purement européennes. Les plateformes que nous préconisons devront interopérer avec les services des Clouds étrangers tout en restant vigilant sur les règles de protection de données..

Impact :

- *Social* : (3) Les plateformes ouvertes et spécialisées dans divers domaines (Industrie 4.0, Smart cities, Banques et Assurance, Santé et bien-être, Mode et beauté, Environnement, Energie...) accompagnées d'une gouvernance agile et de campagnes de communication, aideront à créer une dynamique d'innovation collaborative à tous les niveaux de la société
- *Économique* : (5) Une coopération entre les start-up, les industriels de l'IT traditionnels, et les grands utilisateurs, aux niveaux national et européen, tout en étant ouvert aux solutions Américaines et Asiatiques, permettra à notre écosystème de prendre une part de ce marché qui va toucher tous les secteurs d'activités où une interaction a lieu, tout en minimisant les dépendances vis-à-vis des géants US du secteur.

- *Scientifique* : (4) En fournissant l'accès gratuit à ces plateformes à la communauté scientifique, et en la mettant en contact avec les problèmes réels posés par les Utilisateurs, ne peut qu'aider à l'innovation Utile et le dopage des industriels (fournisseurs et utilisateurs) par la matière grise formée à ces techniques (recrutement, contrats conseils,….)

Recommandations n°2

Recommandation n°2 : Diffuser les méthodes de l'Intelligence Artificielle grâce à une verticalisation sur des thématiques métiers en créant des Hubs IA (« mini cluster » réunissant la recherche, l'enseignement, l'incubation, l'accompagnement et le financement). **(Action PIA3 : AAP Nouveaux écosystèmes d'innovation).**

Résumé : L'enjeu est d'entraîner les algorithmes sur des données métiers pour résoudre le vrai cas d'un vrai client (cf proposition sur la mise à disposition par les grands groupes de leurs données) mais aussi de rendre le résultat de ces algorithmes (qu'ils soient prédictifs ou prescriptifs) utilisables par les entreprises. Pour cela, il faut verticaliser l'usage dans des applications métier qui embarquent les algorithmes en question.

La question de l'IA est essentiellement intégrée en Ile de France, la rendant inaccessible à la plupart des entreprises installées en région. La mise en place de « *corporate centers* » régionaux, sur le modèle allemand et suisse, permettraient de développer l'intelligence collective autour de l'intelligence artificielle. Ces centres favoriseraient les innovations semi-ouvertes et la mise en place d'expertises pointues sur le territoire dans le domaine de l'IA.

Au niveau des financements, et sans changer fondamentalement le paysage de soutien et d'aides, l'Etat pourrait favoriser le financement de projets embarquant de l'IA. Les sociétés proposant des solutions d'IA pourraient être certifiées et répertoriées, pour être éligibles dans ce cadre, à l'échelon régional. Le porteur d'un projet intégrant un prestataire IA, local si possible, pourrait avoir plus de points lors de l'évaluation de son dossier. Au-delà de favoriser l'émergence de projets locaux, valorisés par l'IA, cela permettrait de décentraliser l'IA dans des régions dans lesquelles le tissu actuel n'est pas orienté vers des projets numériques. Un avantage supplémentaire, pour des structures émergentes dans le domaine de l'IA, serait de pouvoir se positionner comme prestataire et donc d'offrir une plus-value sans forcément subir la charge du process de demande d'aides. Cela devrait également permettre de structurer l'IA en filière, idéalement autour d'un ou de champions de l'IA français (acquisition, fusion,…).

Donc pour faire émerger des entreprises qui ont une chance d'être des champions internationaux demain **il faut créer des HUBs IA** (« mini cluster » réunissant la recherche, l'enseignement, l'incubation, l'accompagnement et le financement) :

- Utiliser les filières existantes et structurées en région au sein des pôles de compétitivité et leur donner une dimension IA (comme Mov'eo à Rouen sur Automobile et Mobilité, ou Industrie du Commerce à Lille)
- Coupler cette démarche HUB IA/métier/pôles compétitivité par le développement des centres d'expertise déjà existants en pointe sur le sujet (par exemple Paris Saclay, à l'image de Haïfa en Israël, Bangalore en Inde, Stanford, Cambridge) : tout en privatisant l'unité de valorisation de la recherche en question (aujourd'hui beaucoup trop administrative et déconnectée des marchés).

- Encourager la création (à l'initiative de ces pôles) des équivalents du centre de réalité virtuelle de Laval mais sur chaque métier – rencontre physique entre les grands groupes ou les ETI, les chercheurs et les startups pour travailler ensemble sur des cas concrets. Il est aussi possible de prendre modèle sur le Medialab du MIT au niveau du plateau de Saclay.
- Dédier une partie de l'argent des appels à projet vers l'IA (dédier une part du CNI à cet effet) avec validation par ces HUB IA
- Installer sur ces hubs une relation avec Business France pour permettre de rapidement tester les débouchés des travaux en question à l'étranger
- Associer quelques ETIs ou grands groupes à chaque HUB afin que les données et les besoins/cas d'usage réels soient disponibles et éviter ainsi la répétition des POC (proof of concepts) à faible valeur ajoutée tout en favorisant une forme de R&D collaborative.
- Développer des fonds d'amorçage thématique IA privés, abondés par la BPI

Coût : limité car on se base sur des structures déjà existantes, quelques millions d'euros pour un fonds d'amorçage.

- Constitution et maintien des répertoires des sociétés proposant des services d'IA.
- Intégration dans les grilles d'évaluation des projets d'une rubrique IA
- Co-financement de « corporate centers »
- Création d'un ou de champions français de l'IA.

Estimation de la faisabilité/maturité de la proposition : 4 (structures déjà en places, peut-être des recrutements dédiés à faire).

Priorité : 5 (La France a une avance en Intelligence Artificielle qu'elle doit maintenir)

Impact (gain espéré) : social (4, favoriser le tissu régional de la France en termes de compétitivité et renforcer nos avantages), économique (5, création d'emplois et d'un nouvel écosystème permettant aux grandes entreprises de trouver des solutions à leurs problèmes), scientifique (5, tester en réalité les algorithmes et technologies développées), attractivité (5, scientifiques étrangers qui vont vouloir profiter de cet écosystème favorables à leurs recherches)

Un commentaire libre : Les AAP Ile de France portés par Cap Digital offrent aussi des formats très intéressants comme INNOV'UP Expérimentations, avec des acteurs locaux désirant être terrain d'expérimentation et des process de demande rapides. En ouvrant ce type d'appel au niveau national, en favorisant l'intégration d'IA, cela pourrait favoriser l'émergence de projets sur le territoire.

Un de nos projets, dans le domaine de la simulation automobile, s'est construit autour de notre réseau, en partie cristallisé autour du « *corporate center* » du circuit de Nevers Magny-Cours, l'INKUB. Sans contacts et échanges avec ces partenaires, ils n'auraient jamais embarqués d'IA. Le projet en question est viable sans IA mais prend une autre dimension quand on l'envisage avec de l'IA intégrée. L'idée est vraiment d'ouvrir l'esprit à des personnes qui ont des projets IA compatibles mais n'ont même pas conscience des bénéfices que cela pourraient leur apporter.

Recommandations n°3

Recommandation n°3 : Un programme pour détecter, soutenir et accompagner les talents entrepreneuriaux issus du monde étudiant et universitaire, susceptibles de développer un projet de startup IA.

Le système français de « transfert technologique » est largement structuré autour de la problématique du transfert d'actifs de propriété intellectuelle développés dans les laboratoires publics : c'est en particulier le cas des SATT, dont le modèle économique repose explicitement sur cet aspect, ou encore, pour une large part, du concours iLab.

Cependant, plus largement, le monde étudiant et de la recherche publique est un gisement potentiel de talents susceptibles de développer des projets entrepreneuriaux qui s'appuient sur un savoir-faire qu'ils ont pu développer, indépendamment de l'identification et du transfert d'actifs de propriété intellectuelle. Pour prendre des exemples: Facebook n'a pas été un projet de transfert technologique de l'Université de Harvard, ni SnapChat de l'Université de Stanford.

Dans l'esprit des programmes « French Tech Ticket » ou « French Tech Diversité », le programme reposerait sur 3 actions :

1. détecter ces talents porteurs d'un projet : par exemple un projet d'étudiant de master, d'un jeune ingénieur contractuel dans un laboratoire de recherche, etc. Ni plan d'affaire ni actif de propriété intellectuelle n'est requis ; par exemple le projet peut même n'avoir aucun lien avec l'activité du laboratoire ou de la formation dans lequel travaille le porteur. Par contre l'enjeu est de détecter et faire accréditer le potentiel entrepreneurial du porteur par des tiers eux-mêmes issus de l'écosystème startup : une idée pourrait être que le porteur fasse valoir un « parrain » à son projet, avec un profil d'entrepreneur ou investisseur, qui pourra attester de l'intérêt du projet et des qualités du porteur, tout en s'engageant lui-même à suivre le projet (même a minima) ;

2. financer par une bourse le porteur pour lui permettre de « se lancer » à plein temps pour développer son premier prototype produit, le tester auprès de premiers utilisateurs ou clients potentiels, développer un réseau de soutien autour du projet, etc.; le financement pourra être de durée relativement courte (6 mois ?) mais donner lieu à un livrable identifié et ambitieux (en application des principes de « l'effectuation ») ;

3. dès que le projet démarre, le porteur devra être « immergé » et accompagné dans un environnement entrepreneurial - et non académique, tout en gardant des liens -, d'une part pour bien orienter effectivement le projet vers un projet de startup et d'autre part pour le développement de réseau : il pourra s'agir d'incubateurs, d'accélérateur, etc.

Recommandation n°4

Titre : Mise à disposition d'un « guide d'identification des tâches professionnelles automatisables ».

Résumé : L'intelligence artificielle ayant démontré son potentiel pour remplacer les êtres humains sur certaines tâches (le véhicule autonome étant sans doute l'exemple le plus probant), il semble clair que de plus en plus de tâches seront automatisées. Les métiers eux-mêmes ne seront sans doute pas remplacés du jour au lendemain, mais plutôt progressivement transformés et/ou « augmentés ». Dans ce contexte :

- Les grandes entreprises françaises ont intérêt à identifier les tâches effectuées par leur collaborateur à forte opportunité d'automatisation, afin de réaliser leur transformation par l'achat de solutions ou même l'achat de start-ups (cf. recommandation [Défiscalisation du rachat des startups en IA])

- Les startups françaises en IA ont intérêt à connaître les tâches à forte opportunité d'automatisation des grandes entreprises pour développer des solutions innovantes ou même la création de nouveaux acteurs fondés sur l'automatisation de leurs tâches.

Dans le but d'aider ces acteurs, nous recommandons la création d'un « guide d'identification des tâches professionnelles automatisables ». Celui-ci permettra de faciliter la formalisation des problèmes et catalyser ainsi la mise en place des recommandations [Plateformes Collaboratives de Partage] et [HUBs IA].

Coût : La rédaction du contenue sera réalisée par un groupe d'experts de manière bénévole. Les frais à prévoir concernent la conception graphique, l'impression, la diffusion et la communication, sans doute inférieurs à 100 k€.

Faisabilité / Maturité : 5

La mesure peut être réalisée sans difficulté majeure et dans un délai rapide.

Priorité : 3

Impact : 2 (en accompagnement des autres mesures)

Recommandation n°5

Titre : Etendre au domaine de l'Intelligence Artificielle le Diagnostic Big Data proposé par Bpifrance

Résumé : Bpifrance a lancé fin 2015 le diagnostic Big data afin de permettre aux entreprises (PME et ETI) de disposer d'un diagnostic des enjeux, forces et opportunités dans l'usage de leurs données ou de données disponibles aisément.

Bpifrance propose aux entreprises qui le souhaitent (détection au sein du Réseau Commercial de Bpifrance) de bénéficier d'un contact privilégié avec un expert de l'usage des données massives et offre le financement de 50% du coût de la prestation de conseil dans la limite de 700 €.

Ce diagnostic peut être étendu très aisément aux pratiques de l'Intelligence Artificielle. Ceci nécessite d'identifier des experts susceptibles d'accompagner les entreprises demanderesses.

Coût : Le coût de mise en œuvre est très faible car la plateforme de dépôt est déjà en fonctionnement ; Le coût de la prestation devra être prise sur le budget Bpifrance du P192.

A compléter par une augmentation potentielle des enveloppes budgétaires relatives à ces outils – principalement en subventions et avance récupérables

Faisabilité/Maturité : 5

Impact : 4 ou 5 à chacun des critères (outils existant)

Commentaire : le rapport de France Stratégie démontre une trop grande multiplication des outils de financement. Cette proposition vise à contenir cette inflation en rapprochant le soutien de l'Intelligence Artificielle de celui de l'économie de la donnée.

Recommandation n°6

Titre : Faire des services publics locaux des terrains d'expérimentations avant d'envisager des déploiements nationaux.

Résumé :
L'idée est de favoriser les expérimentations avec les collectivités locales, territoriales ou nationales. Autour de cas concrets (open data, relation avec les usagers, …) permettant d'intégrer de l'IA dans les services publics, nous pourrons non seulement offrir une vitrine aux sociétés travaillant dans le domaine de l'IA mais également fournir des services nouvelles générations performants et facilitant la vie des administrés. En cas de succès, les projets pourraient être portés sur l'ensemble du territoire. Des enveloppes d'accompagnement pourraient être mises en place ou des dispositifs déjà existants pourraient être partiellement réorientés pour financer ce type de projet.

Certains projets à l'échelon local pourraient permettre d'offrir des services à des personnes qui en sont éloignées, agent conversationnel par exemple, pour des raisons de désertification, de manque de personnel ou d'horaires d'ouverture restreints.

L'évaluation d'impact et de coût de la mesure selon un ensemble de critère :

Coût :

- Recensement des besoins
- Détection des sociétés d'IA au niveau des territoires
- Réorientation d'enveloppes de financement pour des accompagnement « appel à des prestataires dans le domaine de l'IA »
- Temps homme pour un référant/une équipe du projet dans l'institution publique.

Estimation de la faisabilité/maturité de la proposition : 5

Priorité : 5

Impact : Social5 ; Economique 3 ; Scientifique 1 ; Attractivité 4.

Commentaire libre (facultatif)

A titre d'exemple Cap Digital a développé un démonstrateur pour les services du cadastre (DGFIP) afin de transformer leur FAQ en point de contact unique interrogeable en langage naturel. L'objectif était de soulager les sollicitations téléphoniques des agents.

Recommandation n°7

Titre : Défiscalisation du rachat de startups en IA

Résumé : Si l'écosystème américain de l'IA contient les GAFAM (Google Apple Facebook Amazon Microsoft), c'est surtout un ensemble dynamique de start-ups dont les GAFAM et autres grands groupes américains sont clients, investisseurs, mais aussi acquéreurs.

- En acquérant des start-ups, les grands groupes intègrent rapidement des technologies innovantes et pointues pour garder leur avantage compétitif.
- Dans ce contexte, les start-ups sont encouragées à développer des technologies innovantes et pointues car même si elles n'arrivent pas en faire des produits vendeurs, elles ont souvent de nombreuses offres de rachat.

La défiscalisation du rachat de startup en IA permettrait de lancer une dynamique similaire au sein de l'écosystème français de l'IA pour à terme atteindre le même niveau d'efficacité.

Coût : Cette mesure pourrait par exemple être intégrée au sein des Crédit Impôt Innovation (CII). Le crédit d'impôt s'appliquerait aux dépenses liées au rachat d'une startup en IA. Le plafond de l'assiette, et éventuellement le taux devront être adaptés.

Faisabilité / Maturité : 3

Les modalités de la défiscalisation doivent encore être définies avec précision (définition d'une « startup en IA », conditions de rachat couvertes, périmètre d'application, …).

Priorité : 5

Cette mesure devrait être mise en place d'autant plus tôt qu'elle prendra du temps à être effective.

Impact : Social : 4 / Economique : 5++ / Scientifique : 3 / Attractivité : 5

Si l'objectif recherché est atteint, l'impact sur l'économie française sera majeur et profond. La dynamique ainsi lancée pourrait faire émerger un ou plusieurs champions français de l'intelligence artificielle, issus des start-ups ou des grandes entreprises historiques françaises, et accroître l'attractivité de la France en tant que pays acquéreur de startups en IA.

Commentaire : La recommandation proposée n'a évidemment pas pour objectif de « forcer » les grands groupes français à acquérir des entreprises « fournisseuses d'IA » sans rapport avec leur cœur de métier. Il s'agit au contraire d'inciter les grands groupes à repenser leur positionnement dans ce nouveau contexte de développement rapide de l'intelligence artificielle. Le positionnement en tant que plateforme semble en particulier être un vecteur de croissance majeur, puisque c'est celui de tous les GAFAM. Apple possède par exemple une expertise dans la conception, le développement et la fabrication de smartphones mais c'est sa position de plateforme qui lui confère son avantage compétitif. Une situation similaire pourrait se présenter dans le cas de l'industrie automobile par exemple, où il est possible que la valeur revienne à terme aux plateformes d'applications pour véhicules connectés / augmentés / autonomes plutôt qu'au constructeurs de ces derniers. Les grands groupes français possèdent de très fortes expertises dans leurs domaines, mais ne disposent a priori ni des compétences ni des technologies nécessaires pour redéfinir leur positionnement en lien avec l'impact de l'intelligence artificielle. Ceux-ci peuvent naturellement les acquérir par des rachats, comme les GAFAM l'ont fait ces dix dernières années.

Recommandation n°8

Titre : Inscrire explicitement la thématique « Intelligence Artificielle » comme priorité des outils d'accompagnement du PIA2 et surtout du PIA 3 en cours de définition

Résumé : Plusieurs dispositifs de financement du PIA2 dans le numérique sont de nature à permettre une réponse à plusieurs des recommandations identifiées à l'issue de cette mission de réflexion sur le développement de l'IA en France. Nous recommandons que ces types de dispositifs, sous cette forme ou autre, soient pérennisés dans le cadre du PIA3 et affichent explicitement une priorité à l'Intelligence Artificielle.

- Favoriser la structuration d'une filière IA en France : le PIAVE (Projets Industriels d'AVEnir) s'inscrivant dans la démarche des 9 solutions de la Nouvelle France Industrielle : alimentation intelligente, confiance numérique, objets intelligents, économie des données, transports de demain, mobilité écologique, ville durable, nouvelles ressources.
  Objectifs : deux types de projets : soit des travaux de développement puis d'industrialisation d'un ou plusieurs produits, procédés ou services non disponibles sur le marché et à fort contenu innovant, soit des travaux visant à renforcer la compétitivité de filières stratégiques françaises. Ces projets doivent démonter un apport concret et déterminant à une filière industrielle et à sa structuration en bénéficiant notamment à plusieurs PME ou ETI issues de cette filière.
  Ce deuxième type de projet parait particulièrement adapté à la structuration d'une filière française IA potentiellement rattaché à la thématique « économie des données « mais clairement transversale à l'ensemble des solutions de la Nouvelle France Industrielle. Ce type de projet permet entre autres de financer **la création d'unités industrielles partagées** permettant à des entreprises d'une même filière s'inscrivant dans une stratégie globale de mutualiser leurs investissements, de participer activement à la stratégie de la filière et de lever certains freins à son développement, la **mise en commun de compétences techniques** (mutualisation les travaux de R&D, la conduite de preuves de concept, tests et autres travaux permettant d'améliorer collectivement leur compétitivité, **mise en place d'outils collaboratifs** (logistique, achats, informatique, intelligence économique, RH, design, marketing, etc.)
  Dans ce cadre, le projet PIAVE peut être porté par une structure fédérant plusieurs entreprises ; le porteur peut associer un laboratoire ou établissement de recherche ou toute structure publique ou privée réalisant ou coordonnant des travaux de R&D. Le financement se fait par un mixte subvention – en avances remboursables. Les projets doivent proposer des programmes d'un montant minimum de 3M€. Le prochain relevé se fera le 28 juillet 2017 à 12h.

  Recommandations pour une relance du dispositif favorisant la diffusion de l'Intelligence Artificielle : Ce type de projet peut permettre de promouvoir et structurer une filière en IA en permettant le déploiement d'infrastructures partagées de R&D, tests, promotion à l'international, compétences, stratégie d'offres, etc. préalablement il faut que le CGI soit sensibilisé aux enjeux de l'Intelligence Artificielle et soit favorable à l'émergence de ce type de projet.

- Favoriser l'intégration de l'IA dans les processus d'innovation des entreprises : le Concours Innovation Numérique (CIN):
  Objectifs : accélérer et amplifier les projets d'innovation de produit et/ou de service, de modèle économique, de marketing, d'usages, de distribution, de design, d'organisation ; créer de la valeur et de la différenciation basées sur de la technologie numérique. L'intelligence Artificielle peut être centrale sur l'ensemble de ces objectifs d'innovation.

Le CIN concerne des projets mono-partenaires ou collaboratifs. A titre indicatif, la taille des projets se situe généralement entre 0.5 à 3 M€ et sont d'une durée de 12 à 18 mois. Le financement se fait 50% en subvention – 50% en avances remboursables.
Le prochain relevé se fera le 5 juillet 2017.
Recommandations pour une relance du dispositif favorisant la diffusion de l'Intelligence Artificielle : L'expérience des relevés passés montre un déficit de compétences dans le domaine de l'IA au sein des entreprises parmi celles qui prétendent s'appuyer sur de l'IA pour mener à bien leur projet d'innovation. Plusieurs projets ont ainsi été écartés de la sélection. Il semble nécessaire de favoriser l'identification et la promotion de compétences dans l'IA qui pourraient être sollicitées par les PME pour optimiser leur démarche d'innovation. Ce constat renvoie également à la nécessité de favoriser une offre de formations professionnelles sur le sujet de l'IA.

- Favoriser la rencontre entre milieu académique et entreprises et le transfert de savoir faire en matière d'IA : Disrupt Campus

  L'appel à projets « Disrupt' Campus » est un dispositif de soutien aux formations à l'innovation numérique et à entrepreneuriat, avec des projets étudiants-entreprises au centre du cursus, financé par le Programme d'Investissement d'Avenir (PIA).

  Un « Disrupt' Campus » (ou campus étudiants-entreprises pour l'innovation de rupture par le numérique) est un cursus de formation à l'entrepreneuriat et à l'innovation numérique en mode « startup », élaboré par un ou plusieurs établissements d'enseignement supérieur (possiblement en association avec un partenaire privé), et associant étroitement des entreprises engagées dans des démarches de transformation numérique.

  Les « projets » sont au cœur de la démarche pédagogique d'un « Disrupt' Campus » : pendant plusieurs semaines les étudiants collaborent avec une équipe d'une entreprise pour résoudre un problème ou de saisir une nouvelle opportunité, ceci grâce au numérique, ses technologies et ses nouveaux modes d'innovation. Un projet doit si possible aboutir à une preuve de concept, typiquement un prototype logiciel ou matériel testable par des utilisateurs ou une adaptation d'un processus de l'entreprise.

- *Objectifs* : **d'une part l'expérimentation de nouveaux cursus d'enseignement supérieur** permettant de mieux préparer les étudiants à l'entrepreneuriat, à l'innovation et à la nouvelle économie du numérique, et **d'autre part la transformation numérique des entreprises françaises.** Au-delà de l'adoption des outils ou de technologies numériques, il s'agit de favoriser l'adoption de nouvelles pratiques de l'innovation (itérations rapides, apprentissage, agilité, souci de l'expérience utilisateur, etc.) et la transition vers de nouveaux modèles économiques (prime à la vitesse, ambition d'emblée internationale, longue traîne, logiques de plateforme, etc.).
- Un projet de « Disrupt' Campus » peut être porté par un ou plusieurs établissements publics d'enseignement supérieur partenaires. Les regroupements d'établissements publics (comme les COMUE par exemple) sont ainsi éligibles. Une entreprise privée peut être aussi partenaire (au côté d'au moins un établissement public) pour apporter soit un savoir-faire et une expérience en animation ou en innovation numérique, soit donner l'accès à ses équipements ou installations.

- Face au succès du premier relevé, une deuxième date sera bientôt communiquée pour un nouveau relevé. Ce dispositif devrait, sous cette forme ou une autre, être reconduit dans le PIA3.

- *Recommandations pour une relance du dispositif favorisant la diffusion de l'Intelligence Artificielle ;* il apparait clairement que ce type de cursus université/entreprise peut permettre de diffuser de la compétence en IA dans l'entreprise sous conditions que les universités intègrent en amont cursus d'ingénieurs en IA à visée opérationnelle.

Recommandation n°9

Recommandation sur la thématique régionale (exemple de l'Occitanie)

Titre : Sur un territoire, fédérer l'ensemble des acteurs concernés par l'IA et le Big Data. Ces acteurs travaillent ensemble afin d'établir une feuille de route ayant pour objectif principal de relever le défi de l'irruption de l'IA et plus, largement, de l'économie des données. L'ambition est renforcer le tissu de la formation et de la recherche et de construire un ensemble de cas d'usage qui seront autant de voies nécessaires à la pérennisation de ces filières

Sous l'égide des pôles Aerospace Valley et Agri Sud-Ouest Innovation ainsi que de l'IRT St Exupéry, les filières d'excellence de la région Occitanie aéronautique, espace, météorologie, agriculture, automobile, IoT ont lancé une initiative visant à fédérer l'ensemble des acteurs concernés par l'IA et le Big Data. Ces acteurs travaillent ensemble afin d'établir une feuille de route ayant pour objectif principal de relever le défi de l'irruption de l'IA et plus, largement, de l'économie des données. L'ambition est renforcer le tissu de la formation et de la recherche et de construire un ensemble de cas d'usage qui seront autant de voies nécessaires à la pérennisation de ces filières. L'enjeu est de leur permettre de bénéficier pleinement du changement de paradigme induit par la révolution numérique. Si chacun d'entre eux a bien pris la mesure du fait numérique, il nous a paru en effet fondamental de bénéficier d'un effet d'échelle, au niveau régional, pour permettre une mise en commun de l'ensemble des besoins nécessaires.

Le groupe de travail « Economie de la donnée » de la région Occitanie formule la proposition que ce type de démarche soit encouragé dans l'ensemble des régions fortes de filières industrielles importantes. L'enracinement local de telles initiatives permettra une expression de besoins en lien avec les acteurs locaux, qu'ils soient issus de la recherche, de la formation mais aussi, des filières industrielles locales qui devront être les principales bénéficiaires de cette démarche. Le PIA3 nous semble un cadre idéal pour faire déboucher ces feuilles de routes localisées.

La région Occitanie ne souhaite pas créer de nouvel objet, qu'il soit un institut ou toute autre forme, ces dispositifs n'étant pas adaptés au mode de fonctionnement de ces technologies par définition ouvertes et dont la réussite est conditionnée par leur capacité à rebondir sans obstacle sur tous les usages de façon transversale.

Des centres de recherche pertinents sur ces domaines existent déjà. Ce qui nous fait défaut c'est l'appropriation par l'aval et la capacité de créer un écosystème opérationnel.

La région Occitanie souhaite s'appuyer sur les structures existantes en les incitant à se regrouper pour adresser au plus près des filières d'usages ces enjeux de développement.

Recommandation n°10

Titre : Construire, en capitalisant sur l'existant, une initiative ambitieuse pour animer l'écosystème IA français et le marketer à l'international afin de positionner la « French Tech » comme un des écosystèmes de référence dans le monde en matière d'IA.

Plusieurs initiatives existent déjà en France pour animer au niveau national et promouvoir à l'international l'écosystème des startups :

- une initiative privée : FranceisAI, qui fédère l'écosystème IA français et est particulièrement active à l'international ;
- un programme public : l'initiative French Tech. Elle n'est pas ciblée spécifiquement sur l'IA, même si « Big Data & AI » est l'une des thématiques qu'elle met en avant à l'international pour promouvoir l'écosystème français.

Afin de soutenir plus spécifiquement des écosystèmes thématiques, l'initiative French Tech a mis en place un dispositif ad hoc, les « réseaux thématiques », qui mobilisent et fédèrent les acteurs concernés (par exemple IoT, Santé, FinTech, etc.). Il n'existe pas de réseau thématique French Tech sur l'IA mais FranceisIA occupe de facto cette fonction : c'est pourquoi la French Tech soutient déjà, au cas par cas, des actions de FranceisAI, via le réseau international des French Tech Hubs et en cofinançant des opérations de communication et d'influence à l'international (rencontre avec des VC internationaux, etc.).

Il s'agirait de construire et soutenir une initiative ambitieuse de mobilisation, de fédération et de marketing international de l'écosystème IA français, en capitalisant sur l'existant : l'initiative French Tech en soutien à FranceisAI. Cette initiative serait dotée de moyens lui permettant de réaliser en propre des opérations (animation et influence internationale) mais aussi de soutenir et accompagner des projets privés qui émergeraient de l'écosystème IA.

Pour l'initiative French Tech cela reviendrait à faire de l'IA une des quelques thématiques sur lesquelles elle focaliserait ses actions et ses outils et concentrerait ses efforts. Pour FranceisAI, ceci lui permettrait de davantage structurer et d'accroitre l'impact et l'envergure de ses actions.

Recommandation n°11

I - Avant-propos

Constat 1, les accélérateurs corporates sont stériles et hermétiques

Ces dernières années, la France a vu émerger de nombreux accélérateurs et lieux d'innovation corporate ces dernières années, issus de la volonté d'attirer des startups et de bâtir des écosystèmes favorisant l'innovation. Cependant les résultats sont en dessous des attentes, et ces lieux n'attirent pas

les entrepreneurs talentueux. Le contre-exemple est les exemples d'accélérateurs corporates comme celui d'Allianz à Nice, qui propose de véritables débouchés pour les startups qui l'intégrent, mettant en place des passerelles significatives avec les différentes business units du groupe et de ses partenaires.

Constat 2, les appels à projet en ligne accélèrent l'innovation

De nombreux dispositifs d'appel à projets en ligne, ou d'appel d'offres public mis en place par des grands groupes, attirent de nombreuses startups désireuses de répondre à ces problématiques et ainsi collaborer avec ces grandes entreprises. Moins coûteux et plus rapides que les accélérateurs corporates, ils permettent d'attirer un plus grand nombre de startups, ingénieurs, chercheurs qui peuvent ainsi plancher sur des problématiques concrètes formulées par les grands groupes. Cette version "en ligne" des accélérateurs corporates est plus agile et permet des résultats concrets, qualitatifs et en plus grande quantité. Ces initiatives prennent souvent la forme de challenges, comme ceux cités ci-dessous:

- Areva Innovation PME http://www.innovationpme.areva.com/

- Le Fab d'Engie https://fab.engie.com/fr/appels-a-projets

- Viva Technology https://challenges.vivatechnology.com/en

- FSMP https://challengedata.ens.fr/en/home

- Cap Digital http://www.capdigital.com/marches/donnee/challenges-big-data/

- Hacker Unit http://www.hackerunit.com/startup

Constat 3, la vision plateforme en opposition avec la vision application

Deux grandes visions s'opposent dans l'innovation corporate, la vision application, qui consiste à développer de nouveaux produits qui répondront aux attentes des consommateurs, toujours plus connectés. C'est par exemple ce que proposent des startups studios corporates comme EHDA d'Euler Hermes, Kamet Ventures d'Axa ou Pathfinder de TheFamily. A cette vision s'oppose celle des plateformes chères à Gilles Babinet, qui invite les grands groupes à se positionner comme des infrastructures en mettant à disposition de l'écosystème des jeux de données. Le plus bel exemple de ce positionnement est celui d'Apple avec l'Apple Store où des millions d'applications ont été développées par des entrepreneurs, makers, ingénieurs et étudiants depuis sa création. Imprédictible et exponentiel, le positionnement d'Apple en tant que plateforme a favorisé l'émergence de startups qui valent aujourd'hui plusieurs milliards. Un autre exemple est celui de GE avec Predix, sa plateforme analytique Cloud pour le milieu industriel. Predix permet ainsi à tous les industriels et acteurs du secteur de l'internet des objets de se plugger sur leur système afin de prévenir les risques de disfonctionnement de leurs installations.

II - Résumé *"IA Startup Tour France"*

Le projet consiste à faire de la France le plus grand accélérateur virtuel jamais créé en France. Pas de lieux physiques où incuber les projets, tout est fait en ligne.

Phase 1 - Appel à projet

Les grands groupes proposent des challenges et remontent des problématiques concrètes issues des besoins des business. A ces problématiques ils ajoutent des jeux de données sous forme d'API. Toutes ces informations sont hébergées sur un portail.

Phase 2 -Candidatures

Les startups, ingénieurs, étudiants, écoles et chercheurs peuvent répondre à ces challenges en formulant leurs solutions directement sur le portail.

Phase 3 - Sélection des projets

Les partenaires, industriels, jury, et citoyens, peuvent tous voter en ligne pour leurs projets favoris et ainsi faire émerger ceux qui seront développés durant la durée du programme. A la manière des campagnes de crowdfunding, le plus grand nombre de personnes est invité à participer à l'émergence des champions de demain.

Phase 4 - Co-Exploration Programme

Les entrepreneurs et industriels travaillent ensemble à la réalisation des projets qui ont été retenus. 6 mois maximum de développement où les grands groupes mettent à disposition des entrepreneurs leurs assets : locaux, ingénieux, données, experts, infrastructures.

Phase 5 - Cérémonies de clôture

Les projets développés seront exposés dans toute la France sur une période de 3 mois: 10/12 villes maximum, 1 par semaine, pour favoriser l'émergence de champions régionaux et permettre des synergies locales startups/industriels. Sur le modèle du "Tour de France", les équipes locales présentent leurs projets à un jury local. Un projet retenu par région.

Phase 6 - Grand messe IA à Paris

Jury final permettant d'exposer les 10 projets retenus en région et ainsi faire émerger le champion IA de demain. L'idée n'est pas de faire émerger une startup d'une région, mais bien une collaboration concrète d'un grand groupe et d'une startup. Le public peut être impliqué dans un "vote du public" et ainsi favoriser la démocratisation de l'intelligence artificielle auprès du plus grand nombre.

III - Coût

Le coût est faible:

- développement d'une plateforme digitale pour héberger les appels à projets

- développement d'un portail pour héberger les données accès sous forme d'API

- développement d'une application mobile pour permettre le vote du public

- communication régionale et nationale

- accueil en région (French Tech)

IV - Faisabilité

Le projet peut être mis en place de manière agile, habile et rapide en utilisant toute la force de frappe de l'écosystème Tech Français:

- La French Tech peut être sollicitée pour permettre le maillage en région

- Les écoles peuvent faire participer leurs étudiants et impliquer leurs professeurs

- Les laboratoires peuvent faire participer leurs chercheurs

- Les fonds d'investissement et les banques peuvent solliciter les entrepreneurs

- Les accélérateurs et incubateurs peuvent relayer les appels à projets

- Les startups peuvent contribuer à la mise en place du dispositif (exemples: Dawex, peut aider à créer le portail de datas. OpenDataSoft, peut aider les grands groupes à mettre leurs données en API)

V - Priorité

Urgent, niveau 5. Il est important de démocratiser l'IA et ainsi favoriser l'émergence de champions Français dans ce domaine applicatif. Les GAFAs acquièrent tous les talents de l'IA.

VI - Impact social

Ce dispositif permet de fédérer l'écosystème autour de problématiques communes, et permet la création d'emplois dans le secteur du digital. Le public étant impliqué, l'IA ne sera plus un secret pour le commun des mortels qui pourra y voir les bénéfices immédiats.

VII - Impact économique

Ce dispositif permettrait aux grands groupes de rattraper leur retard notable sur le digital et ainsi d'éviter leur disparition, qui serait un désastre pour l'économie nationale.

VIII - Impact scientifique

Ce dispositif nous permettrait d'impliquer les écoles, universités, chercheurs et autres groupes de scientifiques afin de travailler ensemble à des cas d'usages et des applications concrètes de l'IA.

IX - Impact attractivité

Ce dispositif a aussi pour vocation d'attirer des entreprises (startups US/UK) en leur offrant des débouchés économiques et de potentiels déploiements. Attractivité des startups mais aussi des grands groupes Européens qui pourraient voir l'opportunité de tester de nouveaux modèles rapides d'exploration de l'IA.

X - Commentaire libre

Faisons de la France le plus grand accélérateur de projets d'Intelligence Artificielle. Nous en sommes capables ! Soyons ambitieux, nous pouvons changer le monde.

Annexe 3 L'Intelligence Artificielle en Occitanie

Automobile : présence de CONTINENTAL

L'évolution des technologies de l'information de la communication et des besoins du marché automobile ont permis l'avènement du concept de véhicule connecté. Le groupe Continental (Domicilié à Toulouse) fort de son expérience d'équipementier automobile électronique, a lancé un projet ambitieux de collecte massive de données à bord des véhicules, ces données devant venir alimenter un cloud où elles seront mélangées à des données de partenaires tiers professionnels de la route ou non. Dans ce concept, les véhicules de toutes marques ont vocation à être connectées dans tous les pays et recevront en retour du cloud des services en vue d'applications d'un niveau de performance sans précédent. Il s'agit donc de milliards de données montantes et descendantes qui seront pour nombre d'entre elles contraintes par le temps réel, la sûreté de fonctionnement, la cyber sécurité et la protection de la vie privée. Face à ces défis, il s'agira de développer des algorithmes d'IA performants permettant d'extraire du big data des services individuels pour chaque véhicule en retour et des services à la collectivité (circulation, secours, sécurité) et aux partenaires tiers.

Spatial et Aéronautique : présence d'Airbus

Le Big Data et l'IA sont au cœur de la stratégie d'Airbus, et d'Airbus Defence & Space. Cela recouvre plusieurs aspects : c'est tout d'abord le chantier de la digitalisation, axé sur la performance opérationnelle et l'intégration des fonctions et de la donnée projet au cours du cycle de vie produit ; cela touche nos cœurs de métiers et nos partenaires industriels. C'est ensuite la valorisation de la donnée produite, et la création de valeur au travers de nouveaux services, développant la richesse mais aussi s'adaptant à des marchés évoluant rapidement et à des business model en perpétuelle redéfinition. C'est enfin une matière unique : celle de la donnée spatiale, imagerie, mais aussi données météorologiques et données d'environnement, recueillie avec une définition, une fréquence, et une précision toujours accrue, ce qui est en soi un défi technique et technologique : comment stocker, comment traiter, comment valoriser un flux toujours accru d'information. C'est une opportunité, la donnée est là et elle permettra de rapprocher le spatial de l'utilisateur. En exploitant habilement les données, via l'IA, les acteurs de l'aéronautique peuvent notamment espérer passer à la maintenance prédictive. Autrement dit, anticiper les pannes afin de limiter les temps d'intervention, donc l'immobilisation des appareils, bête noire des compagnies aériennes et ainsi réduire considérablement les coûts de maintenance des compagnies aériennes. Airbus Helicopters a choisi de développer ses propres modèles mathématiques, et dans cette optique a mis en place une « académie des analytics », pour former ses talents internes. L'accent est également porté sur l'interface. Un des facteurs clés du succès de ces outils réside dans l'expérience utilisateur. Les algorithmes doivent constituer de véritables outils d'aide à la décision et faciliter le suivi des plans d'action. De Constructeur industriel Airbus se réinvente en fournisseur de services, passant d'un modèle de vente à un modèle de location des appareils.

Cette industrie doit donc faire face à un défi en terme de recrutement, d'accès à des technologies développées dans la Silicon Valley mais accessibles à tous, sous réserve de relever le défi et de changer de modèle de développement, l'accès aux communautés Open Source, la révolution du Machine Learning qui modifiera profondément les façons de faire dans la plupart des disciplines de l'ingénierie.

Internet des objets : présence de Sigfox

Le nombre d'objets connectés atteindra 50 milliards à l'horizon 2020, soit plus de 6,5 fois le nombre de dispositifs connectés par personne. Mais où stockerons-nous toutes ces données ? Grâce en grande partie à la promesse de l'Internet des objets, en quelques années seulement, d'ici à 2020, le volume de

données générées dans le monde s'élèvera à 44 zettaoctets, contre 4 zettaoctets à l'heure actuelle. Par conséquent, pour que l'Internet des objets devienne une réalité, nous devons inévitablement répondre dès maintenant au besoin de solutions de stockage et de services transformationnels.

Nous sommes passés du M2M où le barycentre de la valeur était sur l'intégration avec les objets (la communication) à l'IoT où la valeur s'est réorientée vers la donnée. Le challenge à relever est celui de l'équation qui unit collecte, analyse des données et prédictions, nous devons faire face aux défis de la construction de modèles à usage réel et économiquement pertinents, l'analyse prédictive, l'intelligence artificielle, l'exploitation rapide de données massives et fugaces, la sécurisation des données, l'exploitation de données personnelles. Le défi des entreprises de ce secteur (par exemple, SIGFOX entreprise de 325 salariés leader dans l'IOT dont le siège est proche de Toulouse) est de rendre accessibles et traiter de manière simple les problèmes d'usage complexes.

Alors qu'aujourd'hui les entreprises/organisations sont capables d'utiliser des capteurs en tous genres (température, force, pression, géolocalisation, humeur, …) et de les relier à leur système d'information, les quatre principales complexités sont : La gestion du hardware, La gestion du réseau, La collecte des données, Le traitement & la valorisation des données.

Météo : présence du Centre national de recherches météorologiques

Les prévisions météorologiques et l'étude du changement climatique sont indissociables des sciences de la donnée. Les prévisions comme les analyses du climat récent s'appuient sur des données massives collectées par d'innombrables capteurs qui grâce à des traitements numériques (résolution d'équations complexes et analyse statistique) offrent des prévisions ou des estimations du climat de plus en plus précises et fiables.

La demande de prévisions météorologiques est exponentielle et accompagne les progrès de nombreux secteurs dont certains travaillent sur les machines autonomes comme l'automobile, l'aéronautique ou l'agriculture. Les sujets de la météo comme du climat sont en tout point transversal car s'ils utilisent depuis toujours des données in-situ, ils se nourrissent maintenant, largement des données du spatial (plus de 100 millions de données traitées par jour).

Au-delà de ces données "structurées", l'IA aujourd'hui offre un énorme potentiel prometteur en science de l'atmosphère et du climat. Ces données massives devront contribuer à améliorer les prévisions et les interventions d'urgence en cas de catastrophes naturelles. Un sujet d'intérêt récent concerne les données de téléphonie mobile : elles permettaient de mesurer les précipitations au plus près du sol mais aussi de détecter les signaux de stress dans les zones frappées par des catastrophes naturelles.

Face au réchauffement planétaire, la révolution des données est cruciale, car elle :

- aide à visualiser les effets concrets du changement climatique
- favorise le développement de nouvelles solutions pour s'en protéger
- initie de vraies inflexions – tant politiques que comportementales.

Le Centre national de recherches météorologiques (CNRM) est une unité Mixte de Recherche constituée par le **CNRS** et **Météo-France**. Reconnu comme un des leaders mondiaux de la recherche météorologique, le CNRM assure l'essentiel des activités de recherche et coordonne l'ensemble des actions de R&D de Météo-France. Ses activités de recherche sont principalement menées sur les sites de **Toulouse** et de Grenoble.

Les équipes du CNRM s'appuient sur des moyens de haute technologie : supercalculateurs et modèles numériques du système Terre, satellites d'observation, avions instrumentés, stations de mesure fixes ou mobiles, sites instrumentés en montagne, radars profileurs de vent… et coopèrent étroitement avec des laboratoires, universités et structures de recherche français et internationaux.

Agriculture : centre de recherche INRA et pôle de Compétitivité Agri Sud-Ouest Innovation

La 4e révolution agricole sera numérique. Après la mécanisation, les fertilisants et produits de protection des plantes, après les biotechnologies vertes, l'intelligence amenée par la gestion des données agricoles permettra de produire mieux : économiquement, écologiquement et socialement.

Une augmentation de 50% de rendement est atteignable en France grâce à la création de systèmes de production digitalisée. Cela sera nécessaire pour réussir l'un des futurs enjeux mondiaux : nourrir 9 milliards d'habitants à l'horizon 2050. La digitalisation des systèmes de productions permettra une augmentation des rendements en parallèle d'une économie des ressources créatrice de valeur et d'emplois. On peut estimer à une centaine de milliards d'euros le chiffre d'affaires supplémentaire qui est en jeu en France grâce à l'implémentation du numérique dans les secteurs agricoles et agro-alimentaires.

A travers une meilleure exploitation des données, le big data et l'intelligence artificielle auront un impact considérable sur tout le système de production et de transformation des produits agricoles. Ainsi, les agroéquipements intelligents seront au centre de la ferme de demain. Equipés de capteurs, de systèmes embarqués et connectés au cloud, ils pourront être autonomes et agir grâce à des données fournies par un large panel de capteurs (satellites, drones, sur machine, dans le sol). Ils appliqueront la bonne action, au bon moment et au bon endroit. On parle « d'intelligent farming ».

Les premiers projets mis en œuvre par le Pôle de Compétitivité Agri Sud-Ouest Innovation montrent que l'on peut diminuer de plus de moitié l'utilisation de certains pesticides. Il en est de même pour l'utilisation de l'eau. Ainsi, avec des données massives, plus importantes et mieux analysées, l'amélioration de la ferme France peut connaitre une vraie révolution.

Le secteur aval est également concerné : l'agroalimentaire et la valorisation non alimentaire. La digitalisation permettra de mieux répondre aux enjeux de sécurité, de traçabilité. Les entreprises agro-alimentaires pourront formuler de nouveaux produits plus vertueux. Les consommateurs pourront connaitre avec précision la composition et l'origine du produit consommé. De plus, des systèmes intelligents permettront aussi de mieux gérer les flux logistiques et de mieux affecter les ressources amont et les besoins aval sur une maille plus fine, celle du territoire. Cela fera naitre la bioéconomie territoriale.

Dans la recherche, l'exploitation des données permettra aussi de grands progrès en connaissant plus finement les liens entre la génétique, l'environnement et la performance de production des végétaux et des animaux. Les milliards de données stockées, aujourd'hui complexes à analyser, pourraient demain fournir de précieuses informations pour améliorer la génétique des plantes et des animaux.

Enfin, de manière transversale, la maitrise des systèmes de production agricoles et agroalimentaires – les agrochaînes – est complexe : le vivant n'est pas homogène et l'avènement des systèmes d'intelligence artificielle est un moyen supplémentaire pour mieux connaitre ce vivant et en quelque sorte mieux le maitriser : moins de gaspillage, moins d'intrants, plus de rendement. Nourrir les hommes et alimenter l'industrie : telle est l'enjeu de l'agriculture de demain. La complexité de ces systèmes demande une approche nouvelle : celle des données massives et de leur exploitation « intelligente ».

Outre la richesse du territoire dans les domaines agricoles et agroalimentaire, la présence à Toulouse du principal centre de recherche INRA dans le domaine de l'analyse des données agricoles et de

l'intelligence artificielle (Laboratoire MIAT - unité de mathématiques et d'informatique appliquées de Toulouse) constitue un atout majeur sur lequel s'appuyer pour développer des projet big data sur Toulouse et ses environs.

GROUPE DE TRAVAIL 2.3
VEHICULE AUTONOME

Groupe Véhicule autonome

1 - État des lieux

a) Société

La mobilité durable est devenue un enjeu sociétal majeur dans un contexte où l'organisation actuelle des territoires urbains et péri-urbains conduit à une augmentation des échanges. La mobilité est aujourd'hui contrainte par de nombreuses questions, relatives à la sécurité des usagers, à l'accès à l'offre de mobilité, à la dépense énergétique, à l'impact sur l'environnement. Les transports autonomes (véhicule et système d'organisation), peuvent contribuer à déverrouiller certains de ces challenges.

- En 2016, l'ONISR a annoncé le nombre de 3469 morts sur nos routes, chiffre stable par rapport à 2015. 90% des accidents routiers sont causés par des erreurs humaines[1], ainsi une automatisation de la conduite peut contribuer à la réduction du nombre et de la gravité des erreurs de conduite.
- Aujourd'hui, plus de la moitié de la population française n'a pas accès à une offre de transport en commun[2]. Il existe un fort déficit d'offre de service dans les zones peu denses, mais également en périphérie des grandes villes qui bénéficient rarement d'un réseau capillaire de desserte important. Le développement du véhicule autonome, en complément des moyens de transport de masse, offre des opportunités. Par ailleurs, cette inégalité d'accès est encore renforcée pour les plus vulnérables (personnes âgées, handicapés …) qui nécessitent des moyens plus spécifiques et une desserte plus précise jusqu'au dernier kilomètre.
  L'autonomie apporte un changement de vision du véhicule qui, plus qu'un objet, devient un service complémentaire aux transports en commun et de marchandises actuels, notamment en zone urbaine et périurbaine. Ce changement dans les pratiques de mobilité adresse directement des défis sociétaux comme la durabilité, l'éco-responsabilité des systèmes de transport, la diminution des congestions… Il permet également le rapprochement entre les transports individuels et collectifs, et rend possible l'émergence de nouveaux acteurs économiques de transport collectif plus individualisé.
- Enfin les transports en commun eux-mêmes sont également promoteurs de solutions d'autonomie (les premiers métros automatiques datent des années 1980) pour rendre leur exploitation plus flexible. Pour étendre le concept au-delà des métros automatiques, il est aujourd'hui envisagé des transports en commun autonomes sur des sites non dédiés (train autonome ou tramway autonome).

b) Juridique

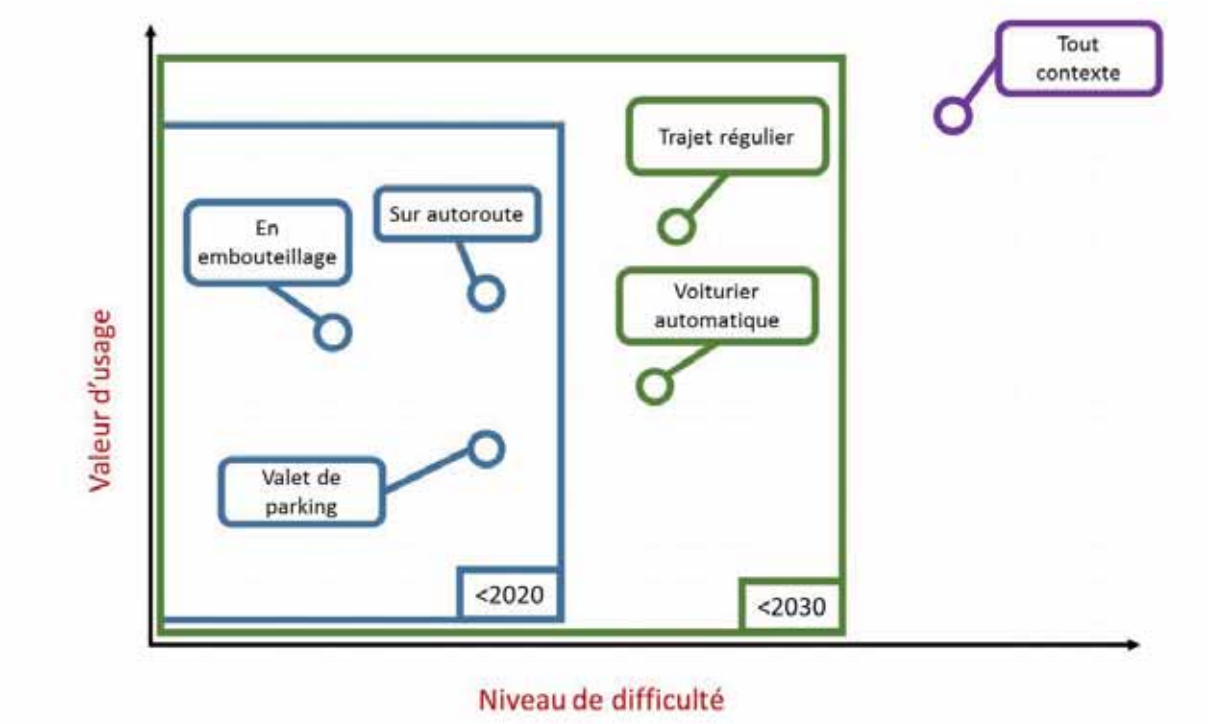
Il existe un enjeu important sur la responsabilité en cas de défaillance du système. Il faudra pouvoir définir le partage de responsabilité entre le concepteur du véhicule, les différents fournisseurs de technologie mais également le conducteur en cas de défaillance ou de mauvais usage du système.

c) Économique et place des industriels français

Depuis fin 2013, et le lancement de la Nouvelle France Industrielle, la France s'est doté d'un programme visant à faciliter le développement technologique et le déploiement du véhicule autonome en France. Ce plan a produit une feuille de route avec un horizon temporel pour le déploiement des différentes applications du véhicule autonome.
Les expérimentations de ces véhicules sont possibles sur routes ouvertes depuis mi-2015 et les principaux acteurs français ont déjà commencé des tests.
Les acteurs français ont un déficit de communication sur ce sujet, cependant les expérimentations en cours de Renault et PSA ainsi que les différents développements de capteurs et d'algorithmes de Valeo montrent que la France reste encore aujourd'hui dans la course.
Par ailleurs, la présence en France de deux constructeurs de navettes autonomes, est un avantage important dans le développement des technologies du véhicule autonome. La mise en place d'une réglementation spécifique pour ces objets permettrait de conforter une place de leader sur le marché des petites navettes de transport en commun autonome.
Dans le cadre de la NFI les acteurs du véhicule autonome ont proposé une feuille de route temporelle en classant les applications du véhicule autonome en fonction de leur arrivée à maturité technologique espérée. Cette feuille de route est similaire temporellement à celle suivie par la plupart des acteurs internationaux crédibles sur le sujet.

---

[1] Association Prévention Routière, Statistique 2012 Principaux Facteurs d'accidents
[2] Ministère des transports, carte UTP 2012



Plan Nouvelle France Industrielle, feuille de route [3]donnant un horizon temporel pour le déploiement des différentes applications du véhicule autonome.

d) Technologique

La France possède des atouts forts, à travers des filières historiques (transport, nucléaire …) sur les technologies liées au véhicule autonome (sûreté de fonctionnement, technologies de capteurs …).
Il existe également un enjeu fort sur le développement de l'intelligence de ces moyens de transport. En effet l'intelligence artificielle est à mobiliser pour la bonne compréhension sémantique de l'environnement du véhicule et pour ses prises de décision mais aussi pour l'optimisation des trajets ou encore la communication avec l'usager.
Ces enjeux font l'objet de la seconde partie de ce rapport, avec pour objectif d'identifier les technologies d'IA nécessaires, à des coûts acceptables, pour favoriser le développement du véhicule autonome en France.

2 - Les enjeux technologiques du véhicule autonome

Cette section a pour objectif de lister les enjeux technologiques du véhicule autonome dans trois domaines concernés par l'intelligence artificielle. Ces différents enjeux sont explicités dans le tableau suivant. En annexe, un tableau détaille les questions prioritaires de recherche (non adressées ou nécessitant un effort de recherche important), surlignées en rouge et mises en lien avec les domaines scientifiques et technologiques de l'IA tels que définis par l'INRIA dans son « livre blanc sur l'IA ».[4]
Ce tableau est décomposé en 3 dimensions distinctes :

- La dimension *véhicule*, qui reprend les thématiques nécessaires au développement de l'intelligence du mobile,
- La dimension *humaine*, qui explicite les thématiques d'intelligence artificielle nécessaire à une meilleure acceptation de ces véhicules,
- La dimension *système*, qui nomme les thématiques à aborder au-delà du véhicule comme par exemple celui des communications entre véhicules ou avec l'infrastructure, la compréhension sémantique des scènes au niveau système, la problématique des données de roulage pour la mise au point des systèmes d'IA...

[3] http://www.pfa-auto.fr/wp-content/uploads/2016/03/Objectifs-de-recherche-Vehicule-Autonome.pdf
[4] https://www.inria.fr/medias/inria/documents/livre-blanc-ia

Dimension Véhicule
Perception et compréhension de l'environnement : Afin de devenir complètement autonome, le véhicule devra être capable de fournir une interprétation correcte des scènes de la circulation automobile. Des problématiques de développement technologiques sont identifiées dans les sujets suivants : - Détection et suivi des autres usagers (taille, position, orientation …) *Notamment, la reconnaissance des gestes et la compréhension de l'intention des usagers.* *Par ailleurs il est important de pouvoir accompagner ces informations d'un score de confiance.* - Détection signalétique (ligne, marquage, panneau…) *La détection de sortie du cas nominale reste problématique. Par exemple aujourd'hui il reste difficile d'identifier les zones de travaux.* - Reconstitution de scène *Afin d'assurer une perception globale d'une scène, il faut pouvoir ordonner et fusionner les différentes informations en une seule représentation cohérente.* - Compréhension des intentions des autres usagers *La reconstitution de scène n'est pas suffisante, il faut également interpréter les comportements récents des différents acteurs de la scène. Il est nécessaire d'anticiper les intentions et motivation des acteurs de la scène.* Par exemple, ne pas dépasser un véhicule qui s'apprête à sortir de l'autoroute.
Localisation : - Localisation, carte locale *Cartographie dynamique garantissant l'autonomie de conduite hors signal pour une mise en situation de sécurité. Perception augmentée, capteur virtuel ou via connaissance collaborative* - Localisation dans Cartographie Haute Définition *Besoin en localisation plus précis que la définition actuelle du GPS, la précision pourrait être amélioré par reconnaissance visuelle de lieu et mise en correspondance d'amers visuels entre caméra et base d'images de la cartographie.* *La production et la mise à jours d'une telle cartographie et couteuse.*
Décision : - Analyse de scénarios complexes *Analyser des scénarios de trafic complexes (avec des véhicules autonomes, non autonomes, des piétons, cyclistes, deux roues…)* - Définition des comportements du véhicule *Afin d'avoir un comportement garantissant la sécurité et acceptable au regard des autres usagers et de la législation, il est nécessaire d'obtenir un compromis entre plusieurs aspect. Ce compromis doit être effectué entre efficacité de la conduite et respect trop formel du code de la route.* Par exemple, le véhicule doit être capable de franchir une ligne blanche pour dépasser un véhicule arrêté bloquant la circulation. - Variantes de décision *Il est primordial pour le confort mais également pour l'acceptabilité du véhicule autonome qu'il puisse adapter sa conduite en fonction du contexte dans lequel est situé le véhicule (localisation, style de conduite ...)*
Contrôle-Commande : - Planification de trajectoire locale *Générer rapidement un ensemble de trajectoires sûres en temps réel avec un environnement incertain, incomplet en zone urbaine et sur route départementale.* - Manœuvres d'urgence *Conduite aux limites de commandabilité et d'adhérence, comportements complexes (dérapages, etc).* - Manœuvres d'anticipation *Positionnement dans les voies pour anticiper la prochaine manoeuvre. Négocier le passage en sécurité dans une voie bloquée. Prévoir des trajectoires quand il n'y a pas de marquage ou pas défini.*
Mise en sécurité, auto-diagnostic *Comprendre la situation et le danger*. Effectuer dans un temps limité et de manière sécuritaire une manœuvre d'urgence avec des capacités du véhicule dégradé (perte d'un capteur, pneu crevé …). Le véhicule doit être capable de connaitre et d'analyser sa situation (par ex une perte de capteur), il doit pouvoir également *reconnaitre qu'il ne sait pas gérer une zone*.

Dimension Humaine
Interaction habitacle Bien qu'autonome, le véhicule peut apporter de nombreux services, relatifs à la conduite, comme la gestion des transitions ou l'aide à la compréhension de l'environnement - Personnalisation : L'assistant personnel est traité en dehors de la problématique véhicule. Il s'agit ici de se focaliser sur les interfaces et les services que l'usager peut avoir en effectuant une convergence des systèmes (assistant personnel et véhicule autonome), comme l'anticipation de détours utiles (modification de l'itinéraire pour passer chez un fleuriste ou tout simplement pour faire le plein) - IHM / ergonomie : il s'agit ici de rendre la conduite compréhensible et prédictible afin de conserver la confiance des utilisateurs, ainsi que de comprendre le fonctionnement du conducteur pour lui présenter les bonnes informations de situations et contexte au moment de la reprise en main. L'objectif est ici de comprendre le comportement classique de conduite et pouvoir apporter des informations pour l'optimiser. - Driver monitoring : même au-delà des phases de transition, il faudra pour un système autonome comprendre le comportement des personnes dans l'habitacle afin de prévoir les comportements inadaptés, comme conduite avec les pieds sur le tableau de bord.
Interaction hors habitacle On distingue dans cette catégorie les interactions avec les autres usagers, qu'ils soient humains (autres conducteurs, usagers vulnérables) ou non (autres véhicules autonomes, gestionnaire d'infrastructure). Les interactions non verbales entre les usagers sont aujourd'hui importantes que ce soit pour la prise de décision (négociation avec un tiers) ou la compréhension des actions (faire comprendre ses intentions). Il faut donc mettre au point, pour gérer ces interactions non-verbales, des algorithmes d'IA : reconnaissance de gestes, actions du VA révélant clairement ses intentions (petits mouvements préalables avant de démarrer réellement à une intersection, …), et vérification de la bonne compréhension des usagers environnants d'après leur réactions.

Dimension Système
Données pour la mise au point des systèmes d'IA pour le Transport - Collecte de données de roulage et Open Data Les données qui servent à concevoir les systèmes autonomes coûtent cher à acquérir et sont vitales pour le développement et la validation des systèmes perceptifs. Pour faciliter et accélérer le déploiement de ce marché, ces données peuvent être mutualisées. L'anonymisation bien pensée couplée au partage des données permettra de disposer de Data Sets annotés avec vérité terrain sur une large échelle. - Critères d'évaluation et base de tests L'enjeu est ici de réussir à établir un catalogue commun à l'échelle de l'Europe des situations de tests aux caractéristiques bien identifiées, nécessaires à l'évaluation des algorithmes de décision, de planification et de contrôle, y compris les capacités d'auto-adaptation et d'apprentissage, et les problématiques éthiques.
Compréhension sémantique des scènes au niveau système - **Supervision (hypervision) des systèmes (de systèmes) de transport autonome.** Les capacités de communication permettent d'envisager un système de supervision et de / routage des véhicule autonomes dans les zones denses ou complexes (une sorte de "Air Traffic Control" automatisée). Il est impératif de bien définir les règles propres au trafic routier incluant la délégation d'autorité. Cette supervision imposera une identification régulière des mobiles vers une infrastructure, une vision collaborative de la supervision et des capacités d'interception. - **Partage d'informations pour l'optimisation et la négociation.** Chaque niveau de perception peut gagner à être partagé avec d'autres véhicules autonomes et avec des infrastructures adaptées. Ex: Envoyer au fournisseur de cartographie que le marquage au sol perçu était anormalement éloigné de la cartographie HD, ou qu'il y a des travaux Ex: Partage d'intentions entre véhicules Le dialogue et la négociation entre véhicules autonomes contribuera à optimiser l'usage du réseau routier en introduisant par exemple des intersections automatisée et « smart ». Un travail important reste à faire au niveau standardisation. - **Hétérogénéité des systèmes autonomes** Plusieurs systèmes autonomes, dans plusieurs versions avec plusieurs modes/capacités de décision pourront partager le même environnement. Dans les algorithmes de décision de trajectoire par exemple, ceux-ci prennent en compte la prévision de la trajectoire / réaction des autres acteurs. Les différentes

versions d'IA impliquées sont des non déterminismes dans les interactions. Ce problème peut être partiellement levé par la communication inter véhicule sous réserve d'interopérabilité. Ex: variabilité des comportements des véhicules (plus ou moins agressifs)
Verrous non fonctionnels pour l'industrialisation - **Explicabilité.** En cas de litige, les technologies devront offrir la possibilité, après rejeu de la situation enregistrée, de faire comprendre à des humains les alternatives choisies ou non et d'expliquer le choix retenu. - **Maintenabilité du système.** Les releases des algorithmes embarqués pourront tirer parti de bases d'apprentissage collectées en mode collaboratif et incrémental. Ceci nécessitera d'optimiser l'information montante et de consolider tous les tests avant mise à jour. En cas de défaillance décisionnelle avérée, elles devront aussi permettre de rapidement corriger, valider et déployer les nouvelles versions sur les véhicules concernés - **Optimisation et Embarquabilité.** Il est évidemment indispensable de faire fonctionner en temps réel et de façon déterministe les algorithmes embarqués de perception et fusion de données, notamment basés sur la vision, et auto-adapter les décisions en fonction du contexte routier. En ingénierie système ceci impose de définir préalablement les standards « d'hébergeabilité » des algorithmes d'IA, de disjoindre harware/software par un partitionnement adéquate. - Cybersécurité - Détection d'attaques Zero-Day. Capacité à travailler sur des données cryptées (homomorphique)

3 - Recommandation

1 - Promouvoir une approche logique complémentaire à l'apprentissage pour la prise de décisions.

Outre la perception, la prise de décisions durant la conduite est un enjeu crucial pour les voitures autonomes. Son but est, tout en garantissant la sécurité, de marier les objectifs parfois contradictoires que sont le respect de l'itinéraire, le respect du code de la route et la qualité de conduite. Il semble crucial d'encourager les approches respectant ces qualités:

- La capacité à justifier du raisonnement qui a été tenu en cas de litige (expliquabilité) ;
- La capacité à modifier/corriger pour créer facilement des variantes de décision (contrôlabilité).

L'approche logique à base de règles (ou autre approche algorithmique logique) a le double avantage d'être une expertise française, et de ne pas être dépendante de données d'apprentissage.
Cette approche logique ne s'oppose pas à l'approche par apprentissage qui permet de modéliser des comportements que l'on ne peut appréhender que par l'observation et l'apprentissage (comportement dans les ronds-points par exemple).

2 - Préciser le code de la route afin de fournir une interprétation sans ambigüité pour le véhicule autonome.

Le code contient de nombreuses ambigüités qui demandent au conducteur d'interpréter le code de la route en fonction de la situation qu'il observe. Afin que tous les véhicules autonomes respectent et interprètent le code de la route d'une manière identique, il est nécessaire de lever ces ambigüités. Une solution serait que l'État fournisse une interprétation claire des situations ambiguës du code de la route. Par exemple, en cas de priorité à droite ou d'entrée sur un rond-point, jusqu'à quelle distance (spatiale et/ou temporelle) du véhicule ayant priorité doit-on réellement la lui laisser ? Ou encore, dans le code de la route il est interdit de rouler sur le trottoir, sauf en cas d'urgence, mais on ne définit jamais cette urgence.

3 - Mutualisation des bases de données d'enregistrements capteurs (caméra, radar, GPS, accéléromètre, etc…) pour la compréhension, l'apprentissage statistique et l'évaluation.

De manière conjuguée à l'augmentation de la puissance de calcul, l'accès à de grandes bases de données a permis ces derniers temps de très fortes progressions dans le domaine de l'apprentissage. Ces bases de données servent aussi à la compréhension des phénomènes et à l'évaluation des algorithmes développés. Or, elles ne sont pas adaptées à la problématique Véhicule Autonome.
Il s'agit dans cette action de développer, mutualiser et entretenir des bases de données annotées d'une taille conséquente, variée et représentatives des environnements routiers et des conditions de circulation sur des réseaux Français et Européens, assez différents de ceux Nord-Américains. Ces bases de données doivent agréger différentes modalités de perception. Il faut aussi en faciliter la mise en place et en favoriser l'accès que ce soit pour la recherche privée et publique. A terme, ces bases pourraient servir à la communauté scientifique internationale pour l'évaluation des algorithmes.

4 - Création d'une base de données européenne et des méthodes permettant l'enrichissement des cartographies haute définition des données de sécurité

Synthèse : La circulation des véhicules autonomes dans des milieux complexes (par exemple en centre-ville) nécessite la mise à disposition de l'intelligence d'une cartographie haute définition contenant de nombreuses informations. Certaines informations possèdent un caractère critique pour le fonctionnement du véhicule autonome de manière sécuritaire (positionnement des zones de travaux, accidents, météo ...). Ainsi la mise en place au niveau européen d'une plateforme d'échange de ces données semble critique à terme pour assurer le fonctionnement de manière sécurisé du véhicule autonome. Il faut de plus développer des méthodes de mises à jour automatiques de ces bases, fondées sur des informations incertaines et limitées. Afin d'éviter tout problème de concurrence, cette plateforme pourrait être opéré par la puissance publique ou par une entreprise avec une DSP.

5 - Pour l'écriture de la réglementation internationale, porter une position proposant une redondance algorithmique dans les véhicules et développer le positionnement des différents champs thématiques de l'IA dans le Véhicule Autonome.

Synthèse : La redondance algorithmique qui est courante en aéronautique, où plusieurs composants confirment un résultat, doit être appliquée à la voiture autonome pour les décisions jugées critiques, comme le fait de mettre la voiture en sûreté dans la bande d'arrêt d'urgence, ou d'enfreindre certains aspects du code de la route pour des raisons de sécurité.. Ces décisions doivent confronter l'avis de divers systèmes décisionnels (construits par différentes équipes et utilisant des technologies/approches différentes). Ainsi, l'intelligence Artificielle peut intervenir dans différents domaines du VA, il convient d'en explorer les capacités et l'adéquation avec les différentes fonctions constituantes d'un VA.

Annexe 1

Paul Labrogere	Directeur Programme - Transport Autonome (IRT SystemX)
Thibaut Ferreira	chargé de mission (DGE - SI/Bureau de l'industrie automobile)
Patrice Aknin	Directeur scientifique (IRT SystemX)
Sebastien Glaser	Responsable du projet véhicule à conduite déléguée (ITE Vedecom)
Philippe Morignot	Chercheur senior (ITE Vedecom)
Claude Le Tallec	ONERA
William Ledoux	Spirops
Jerôme Hoboian	Spirops
Huges Cheritel	SNCF
Emmanuel Manceau	Quantmetry
Christian Chaumette	Alstom
Vanessa Picron	Valeo
Fabien Moutarde	Mines Paris Tech
Fawzi Nashashib	INRIA
Olivier Trebucq	INRIA
Jean-François Sencerin	Renault
Vincent Abadie	PSA
Frederic Large	PSA

Annexe 2

Ce tableau détaille les questions prioritaires de recherche (non adressées ou nécessitant un effort de recherche important), surlignées en rouge et mises en lien avec les domaines scientifiques et technologiques de l'IA.

	Domaine
Dimension Véhicule	
Perception et compréhension de l'environnement	
Détection et suivi des autres usagers (taille, position, orientation, vitesse, accélération, type, ...) Certains sujets comme la détection et catégorisation d'objets vus par une caméra sont bien avancés, mais d'autres sont plus lointains, par exemple la catégorisation d'objets vus par un Laser basse résolution ou la détection précise de piétons quel que soit le capteur. Reconnaissance des gestes et compréhension de l'intention (introspection des piétons et des conducteurs humains). Contextualisation et sémantisation. La redondance est néanmoins importante car tous les capteurs ont des points faibles sous certaines conditions (meteo, luminosité, …) Il est important que ces informations soient accompagnées d'un score de confiance, Cf validation du système	Traitement du signal, Image Apprentissage statistique Base connaissance, Inférence
Détection signalétique (lignes, marquages, panneaux, barrières, feux, ...) Hormis les barrières (détectables par le Lidar) on se repose souvent surtout sur des caméras pour percevoir ce type d'informations, ce qui n'offre pas une grande redondance. La détection par vision des feux de signalisation reste encore à fiabiliser. Ainsi que la sortie du cas nominal (zones de travaux…).	Analyse d'image, Traitement du signal, Apprentissage statistique
Reconstitution de scène Les différents capteurs perçoivent sous des formes différentes (images, points, lignes, écho…) avec des horloges, fréquences, temps de traitement et biais différents. Il faut donc fusionner tout ça en une seule représentation cohérente de la scène autour du véhicule à l'instant présent, éventuellement en interaction avec la cartographie HD si elle est présente.	Traitement du signal, Suivi d'objet, Cohérence spatio-temporelles, Gestion de l'incertitude Apprentissage auto ?
Compréhension de scène La reconstitution de scène ne suffit pas à résoudre toutes les problématiques décisionnelles de conduite pour des niveaux élevés d'autonomie. Il faut en effet être capable de transformer les comportements récents des différents acteurs de la scène, éventuellement avec l'aide de la cartographie HD, en anticipations des intentions et motivations des véhicules autour, et pas seulement sous la forme de trajectoires. Ex: pour ne pas se rabattre pour laisser un véhicule nous dépasser alors qu'il s'apprête à nous doubler par la droite. Ex: ne pas dépasser inutilement un véhicule qui s'apprête à sortir ou qui vient de s'insérer et va accélérer.	Analyse d'image Traitement du signal Apprentissage auto Collection de data
Localisation	
Localisation, carte locale Cartographie dynamique garantissant l'autonomie de conduite hors signal pour une mise en situation de sécurité. Perception augmentée, capteur virtuel ou via	Traitement du signal Connaissance

		connaissance collaborative.	
		Localisation dans Cartographie Haute Définition Une cartographie HD intègre les détails géométriques de chaque voie de circulation avec une précision relative de quelques centimètres, ainsi que les types de marquages, panneaux, barrières etc.. Elle permet de redonder la détection de signalétique, de calculer des itinéraires précis, et contribue à la compréhension globale de la scène. Précision relative en zone urbaine. Besoin de localisation + précise que GPS, par reconnaissance visuelle de lieu et mise en correspondance amers visuels entre caméras et base d'images de la cartographie. La production et la mise à jour d'une telle cartographie HD est coûteuse. Les formats prototypes sont en développement ainsi que les premières acquisitions sur le sol Français.	Traitement du signal Apprentissage automatique
	Décision		
		Analyse scénarios complexes Analyser des scénarios de trafic complexes (avec des véhicules autonomes, non autonomes, des piétons, cyclistes, deux roues…) Intersection, rond point, dépassement sur nationale	
		Définition des comportements du véhicule Développer et évaluer des méthodes de négociation et de décision, des algorithmes de planification de chemins, pour planifier une trajectoire prenant en compte l'évolution des mobiles perçus dans l'environnement, les contraintes de risque, de consommation et de confort ainsi que les usages locaux (code de la route …) Dès lors que le véhicule peut décider seul de changer de voie, il doit faire cohabiter: 1. La sécurité des passagers et des autres usagers de la route qui doit être garantie 2. Le confort de conduite (ne pas dépasser incessamment, ne pas bloquer les véhicules rapides) **3.** L'efficacité (ne pas non plus rester bloqué derrière un camion) 4. Relaxation du code de la route (dépassement par la droite) 5. Le respect de l'itinéraire (savoir s'insérer pour ne pas rater une sortie) Ces quatres derniers aspects peuvent être mutuellement incompatibles. L'enjeu principal est ici l'obtention d'un compromis **garantissant la sécurité** et acceptable au regard des autres occupants, des autres véhicules et de la législation. Cet environnement accentue les contraintes de qualité sur les perceptions: 1. Pour faire un dépassement: une meilleure gestion des risques/probabilités de présence d'un véhicule 2. Pour faire face à un véhicule en train de dépasser sur notre voie: une bonne compréhension des intentions de chaque véhicule et des autres usagers La conduite en ville ajoute les difficultés suivantes: 1. Gestion des piétons (anticipation de comportement piéton) 2. Stratégies d'anticipation de flux: a. Ex: Ne pas s'engager si on risque de rester bloquer et de géner. b. Ex: Ne pas s'engager dans une voie avec certains véhicules lent c. Ex: Ne pas se placer derrière des véhicules qui vont à gauche si on va tout droit, pour éviter de rester bloqué 3. Gestion du compromis efficacité vs respect trop formel du code de la route (souvent ambigüe) : eg savoir franchir ligne continue pour dépasser véhicule arrêté bloquant (livraisons…), savoir "forcer" un peu le passage pour s'engager sur un rond-point, ou franchir une intersection avec priorité à droite (jusqu'à quelle distance appliquer la priorité aux véhicules déjà sur rond-point ou bien la priorité à droite ?... 4. Comportements et manoeuvres extraordinaires . Véhicules prioritaires a. Véhicules en double file	

	Variantes de décision Que ce soit pour adapter la conduite aux différents pays, pour insuffler une image de marque dans la conduite, ou bien enfin pour permettre à l'utilisateur d'ajuster au cas par cas le style de conduite (Eco, Confort, Pressé) désiré, les technologies devront offrir différents niveaux de contrôle (réglages, ajustement de règles, ...) Identification du style de conduite du conducteur (souple, strict, …)	Algorithmique et Aide à la décision Apprentissage automatique
	Explicabilité, En cas de litige, les technologies devront offrir la possibilité, après un rejeu de la situation enregistrée, de comprendre les alternatives choisies ou non et d'expliquer le choix qui a été fait.	Algorithmique et Aide à la décision Apprentissage automatique
Contrôle-Commande		
	Planification (court-terme) de trajectoire locale Recherche d'un chemin et suivi de ce chemin Générer rapidement un ensemble de trajectoires sûres en temps réel avec un environnement incertain, incomplet en zone urbaine et sur route départementale.	Robotique
	Manoeuvres d'urgence Conduite aux limites de commandabilité et d'adhérence, comportements complexes (dérapages, etc).	Robotique Apprentissage automatique
	Apprentissage du comportement de conduite du conducteur Vision microscopique (et non macroscopique), appliquer le style de conduite du conducteur	
	Manoeuvres d'anticipation Positionnement dans les voies pour anticiper la prochaine manoeuvre. Négocier le passage en sécurité dans une voie bloquée (faire un petit avancement pour voir la réaction des autres) Prévoir des trajectoires quand il n'y a pas de marquage ou pas défini. Créer une voie virtuelle lorsqu'un véhicule prioritaire (ambulance) arrive ou lorsqu'un véhicule 2 roues arrive (par ex., conduite sur le périphérique lorsqu'une moto dépasse).	
Mise en sécurité, auto-diagnostic		
	Compréhension de la situation et du danger. Se ranger sur la bande d'arrêt d'urgence (perte d'un capteur ou pneu crevé, puis clignotant, puis ranger). Moins de contrainte de temps que les manoeuvres d'urgence. Capacité du système à reconnaître de l'information sur lui même (par ex., pneu qui éclate). Comprend véhicule et conducteur (driver monitoring) et environnement (le véhicule peut analyser qu'il ne sait pas gérer une zone), réaction du véhicule qui détecte un malaise du conducteur (se mettre en sécurité et appeler les secours, conduire à l'hôpital).	Algorithmique et aide à la décision. Connaissance.

Dimension Humaine		
Interactions Habitacle		
Assistant Personnel Une grosse partie de l'assistant personnel ne dépend pas du véhicule. On se focalise ici sur la personnalisation, adaptation aux habitudes de l'usager qui permet de faire oublier la technologie. Quelle standardisation des échanges (contenu) pour interfacer les véhicules et les équipements (smartphones) des usagers? Par exemple: Anticipation de trajets et de détours utiles (faire le plein), proposer à une navigation pour s'adapter à l'agenda de l'utilisateur (passer chez le fleuriste).	Apprentissage Automatique	
Confiance des utilisateurs Rendre compréhensible et anticipable les réactions et décisions du véhicule autonome pour conserver la confiance de l'utilisateur dans l'autonomie du véhicule	Apprentissage Automatique	
Phase de transition Garantir la reprise en sécurité de la conduite par un humain. Comment présenter l'information avec une personnalisation adaptée?		
Monitoring de l'habitacle Surveillance de l'état du conducteur pour comprendre son activité et sa capacité à reprendre la main. Détection de comportement inapproprié pour le respect de la sécurité et de la protection des usagers..	Apprentissage Automatique	
Personnalisation Profilage de l'usage pour faciliter l'auto-partage	Apprentissage automatique	
Interactions Hors Habitacle		
Interactions avec les autres usagers de la routes Les interactions sont de trois natures: 1. Faire comprendre ses intentions ("allez-y, traversez", par communication visuelle, sonore et/ou comportement dynamique véhicule) 2. Négocier avec un tiers 3. Passage en mode prioritaire	Apprentissage automatique	

Dimension système		
Données pour la mise au point des systèmes d'IA pour le Transport		
Collecte de données capteurs et OpenData Les données qui servent à concevoir les systèmes autonomes coûtent cher à acquérir et sont vitales pour le développement et la validations des systèmes perceptifs. Or ces données peuvent être mutualisées pour avancer plus vite et faciliter l'accès à ce marché. Anonymisation et partage des données. Datasets annoté et vérité terrain à large échelle.		
Critères d'évaluation et base de tests Etablir un catalogue commun à l'échelle de l'Europe des caractéristiques nécessaires des algorithmes de décision, de planification et de contrôle, y compris les capacités d'auto-adaptation et d'apprentissage, et les problématiques éthiques. Styles et habitudes et conduite en fonction de la région ou du pays (par ex. conduite à gauche en Angleterre, à droite en France, piétons ultra-prioritaires en Suisse/Allemagne, pas tellement à Paris...). Permis de conduire du véhicule.		
Echanges de données instantané & compréhension sémantique des scènes au niveau système		
Supervision (hypervision) des systèmes (de systèmes) de transport autonome. On peut envisager un système de supervision et de / routage des véhicule autonomes dans les zones denses ou complexe (une sorte de "Air Traffic Control" automatisée). Définition des règles propres au trafic routier, la délégation d'autorité. Identification régulière vers une infrastructure, aspect collaboratif de la supervision et capacité d'interception?	Algorithmique et aide à la décision	
Partage d'informations pour l'optimisation et négociation Chaque niveau de perception peut gagner à être partagé avec d'autres véhicules autonomes et avec des infrastructures adaptées. Ex: Envoyer au fournisseur de cartographie que les lignes perçues était anormalement éloignées de la cartographie HD, ou qu'il y a des travaux Ex: Partage d'intentions entre véhicules Dialogue et négociation entre véhicules autonomes pour l'optimisation de l'usage du réseau routier. Intersection automatisée Standardisation et accord des parties		
Hétérogénéité des systèmes autonomes Plusieurs systèmes autonomes, dans plusieurs versions avec plusieurs mode/capacité de décision pourrons partager le même environnement. Dans les algorithme de décision de trajectoire par exemple, ceux-ci prennent en compte la prévision qu'ils font de la trajectoire / réaction des autres acteurs. Les différentes versions d'IA impliquées sont une source supplémentaire de non déterminisme dans les interactions. Ce problème peut-être partiellement levé par la communication partiellement levées par la communication inter véhicule sous réserve d'interopérabilité. Par exemple: variabilité des comportements des véhicules (plus ou moins agressifs)		
Verrous non fonctionnels pour l'industrialisation		
Explicabilité En cas de litige, les technologies devront offrir la possibilité, après un rejeu de la situation enregistrée, de faire comprendre à des humains les alternatives choisies ou non et d'expliquer le choix qui a été fait.	Algorithmique et Aide à la décision Apprentissage Automatique	

		Maintenabilité du système. Constitution de bases d'apprentissage. Une flotte fait remonter de l'information et la redescend pour effectuer des mises à jour. Optimisation pour trier l'information remontante.	Algorithmique et Aide à la décision Apprentissage Automatique	
		Mise à jour et Validation En cas de défaillance décisionnelle avérée, elle devront aussi permettre de rapidement corriger, valider et déployer sur les véhicules concernés	Algorithmique et Aide à la décision Apprentissage Automatique	
		Optimisation et Embarquabilité Faire fonctionner en temps réel et de façon déterministe les algorithmes embarqués de perception et fusion de données, notamment basés sur la vision, et auto-adapter les décisions en fonction du contexte routier. Définir des standards d'hébergeabilité des algorithmes d'IA, capacité d'abstraire le matériel.	Algorithmique et Aide à la décision Apprentissage Automatique	
		Cybersécurité Détection d'attaques Zero-Day. Capactié à travailler sur des données cryptées (homomorphique)		

GROUPE DE TRAVAIL 2.4
RELATION CLIENT

Groupe de travail Intelligence artificielle et relation client

Synthèse finale - état des lieux, enjeux et recommandations

Périmètre d'étude du groupe de travail

La gestion de la **relation client** regroupe l'ensemble des dispositifs ou opérations de marketing et de support ayant pour but d'optimiser la qualité de l'**expérience client**, d'attirer et de fidéliser le consommateur, et de maximiser le chiffre d'affaires ou la marge par client. La relation client englobe la totalité des **interactions**, à **tous les stades du parcours** d'achat puis d'usage d'un bien ou d'un service, et au travers de **tous les canaux de contacts** disponibles. Sa maîtrise constitue un facteur déterminant de la compétitivité hors coût d'une entreprise. Connaître finement chaque consommateur, le séduire efficacement, le servir au mieux, et ainsi le satisfaire pour le retenir, tels sont les objectifs de ce champ d'actions multiforme qu'est la gestion de la relation client.

« Selon les projections du cabinet Gartner, en 2020, 85 % des interactions avec les entreprises se passeront d'êtres humains. Selon la même source, en 2018, 20 % des contenus rédactionnels émis par les marques seront rédigés par des machines, lesquelles contrôleront aussi 40 % des interactions mobiles. Les agents virtuels pourront parfaitement traiter les demandes les plus basiques, standardisées, quand les êtres humains, eux, se concentreront sur les demandes les plus spécifiques. » (Relationclientmag.fr).

C'est à cet essor de l'intelligence artificielle dans le champ de la relation client, à ses conséquences, aux opportunités et menaces que ce phénomène suscite, et aux mesures d'accompagnement de nature à l'orienter dans un sens favorable pour l'économie et la société, que ce groupe de travail s'est consacré.

Etat des lieux

La **transformation numérique** de l'économie affecte en profondeur l'organisation et le fonctionnement des entreprises, et les processus constitutifs de la relation client n'échappent pas à cette mutation.
Depuis l'époque où l'achat d'un produit se faisait uniquement au téléphone après avoir écouté ou vu une publicité, ou en boutique et avec l'aide d'un vendeur, les modes de consommation et de distribution ont bien changé. De nouvelles technologies et des concepts commerciaux et marketing novateurs ont fait leur apparition, révolutionnant les habitudes des consommateurs et leurs attentes, comme les stratégies de segmentation, de promotion, de démarchage et de fidélisation des entreprises. Le client d'aujourd'hui est « **connecté et multicanal** », et l'échange d'informations prend une place prépondérante dans l'acquisition puis la jouissance des biens et services destinés à satisfaire ses besoins ou ses envies. Après la généralisation de l'accès fixe à internet, c'est depuis plusieurs années la diffusion des smartphones et tablettes qui a pris le relais en tant qu'agent de cette révolution client.

La **multiplication des interactions**, en amont, au moment, et en aval de l'acte d'achat, se traduit par un accroissement sans précédent de la **quantité de données** susceptibles d'être recueillies, pour chaque consommateur, sur son historique de consultation et de commande, sur ses avis postés sur un article ou une enseigne, sur son comportement de navigation, sur ses centres d'intérêt, sur son profil sociologique, et autres informations pertinentes.

L'enjeu pour les marques est de construire une **relation attentionnée** avec leurs clients, et ce, malgré la distance et l'automatisation introduites par l'irruption du numérique ; l'empathie, la bienveillance, la personnalisation demeurent des valeurs premières pour susciter les émotions qui convaincront, rassureront, satisferont et fidéliseront un client.

L'intelligence artificielle est entrée depuis plusieurs années dans l'univers de la gestion de la relation client. Son utilisation s'observe, schématiquement, à deux niveaux, qui pourraient être qualifiés de back office et de front line, et gagnent à s'articuler étroitement : l'**analyse des données** et la

détermination des actions appropriées d'une part (marketing prédictif et personnalisé, au besoin en temps réel), les **interactions tout le long du parcours client** d'autre part (relation dématérialisée ou physique automatisée et contextualisée). Les applications de l'intelligence artificielle à la gestion de la relation client se situent, pour certaines, au stade de l'émergence et des premiers prototypes, pour d'autres, en phase de montée en puissance et de déploiement à grande échelle.

Face à la forte augmentation des données disponibles, l'accroissement des capacités de traitement et les progrès des algorithmes d'aide à la décision rendent possible une connaissance et une compréhension accrues des consommateurs, autorisant une ultra-personnalisation de la relation client. C'est l'ère du **commerce cognitif**, ne faisant plus simplement appel à des statistiques avancées, mais recourant à des procédés évolués de détection des besoins et désirs de chaque client, d'extraction de motifs et tendances, et de déduction, par raisonnement simulé, des réponses adéquates.
Il devient possible d'adapter en temps réel la communication avec le consommateur, en fonction de son attitude et de ses réactions. Les deux couches où se déploie l'intelligence artificielle s'adossent ainsi tour à tour l'une à l'autre : en croisant des données de nature complémentaire, des **recommandations** plus opportunes peuvent être élaborées, et proposées en retour, au moyen d'un **assistant shopping digital** ou, mieux encore, sous la forme d'une **boutique virtuelle** entièrement sur mesure. L'objectif est de **capter** une proportion plus élevée de prospects et d'**augmenter le taux de transformation** de visites en commandes ; il s'agit aussi de **fidéliser** les clients déjà en portefeuille.
Les recommandations peuvent ne pas être directement présentées au client, mais être transmises aux conseillers pour les aider à préparer au mieux leurs entretiens en sélectionnant les propositions commerciales à privilégier, ou pour déterminer périodiquement les clients à contacter en priorité.
Les enseignements tirés de l'**analyse textuelle** de commentaires clients peuvent en outre être valorisés par les autres départements (développement, production, qualité, marketing…).
La **publicité ciblée**, réduisant les sollicitations improductives, s'enrichit également de ces techniques. L'intelligence artificielle fait passer d'un consumérisme passif à une **fabrique collective** centrée sur le client et le faire soi-même.

Un autre cas d'usage de l'intelligence artificielle dans le champ de la relation client recouvre les **agents virtuels intelligents**, capables de simuler une conversation, plus ou moins élaborée, avec l'utilisateur, par écrit ou à l'oral, en vue de lui apporter un service dans une large variété de situations. Ces outils se prêtent aussi à un usage interne, en direction par exemple du personnel. À la différence des chatbots traditionnels, les agents conversationnels intelligents sont en mesure de traiter des requêtes en langage naturel, peu normées voire mal formulées, et dont la signification est dépendante du contexte. Les arguments en faveur du recours aux agents virtuels intelligents sont multiples : **immédiateté** des échanges et **continuité** du service, **réduction des coûts** en dégageant les téléconseillers des tâches routinières à faible valeur ajoutée et en les **réaffectant** à des missions plus qualitatives, possibilité de traiter ainsi davantage de sollicitations en parallèle, gain d'indépendance pour l'usager qui apprécie l'**autonomie assistée**…

Les **robots d'interaction** font leur apparition parmi la panoplie de procédés mis en œuvre pour améliorer l'expérience client/usager dans les lieux accueillant physiquement du public. Ils interviennent comme un maillon dans la chaîne de service en déchargeant le personnel des tâches simples et répétitives, ou en introduisant des prestations additionnelles. Les capacités des robots disponibles à un tarif accessible sont encore très réduites, en matière de compréhension du langage parlé notamment, ce qui oblige souvent les concepteurs à prévoir un autre canal de communication, comme une tablette tactile.

Des technologies de reconnaissance vocale, de compréhension et d'interprétation du langage naturel, ou encore de gestion de dialogue doivent, dans les deux cas, être mobilisées. Le bénéfice attendu des développements de l'intelligence artificielle porte sur l'amélioration de la qualité et de l'étendue des interactions, en vue d'un recours plus fluide et intuitif aux services. Un tel progrès passe aussi par une meilleure prise en compte des humeurs et émotions manifestées par l'utilisateur, et par une capacité renforcée d'adaptation au contexte. La reconnaissance faciale fait également partie des aptitudes requises en ce qui concerne les robots physiques.

Les entreprises sont appelées à repenser, en l'envisageant sous toutes ses facettes, leur **management de la relation client**, et au premier chef leur politique de gestion de la donnée ainsi que leur

appréhension du parcours client, qui se doit d'être saisi dans son intégralité. Il convient de ne pas négliger les risques qu'elles encourent face à la **domination**, de plus en plus accentuée, de quelques acteurs, pour l'essentiel étrangers, en matière de concentration des flux d'échanges et de collecte de la donnée. Préalablement au déploiement de tout outil faisant appel à l'intelligence artificielle, un processus de **conduite du changement** doit être engagé en amont pour procéder aux réaménagements organisationnels et humains nécessaires. Pour être à même de garantir une expérience client positive et fructueuse, apparaissent déterminants une parfaite intégration des solutions au système d'information, le décloisonnement des composantes verticales, la possibilité d'un accès multicanal et en même temps cohérent pour l'utilisateur n'omettant aucune des passerelles les plus prisées, et le design et l'ergonomie des interfaces autorisant une expression intuitive du besoin et des échanges sur un mode conversationnel.

La mise au point et la commercialisation de solutions offrant des fonctionnalités répondant pleinement aux besoins des clients (organisation ou consommateur) supposent une capacité opérationnelle à développer des **outils transversaux** mobilisant des **briques d'intelligence artificielle** complémentaires, mais aussi, le cas échéant, des composantes de tout autre ordre (ex : mécatronique, design, son…). La force d'une entreprise peut résider dans une spécialisation reconnue dans une ou plusieurs technologies avancées, ou dans ce savoir-faire en matière d'imbrication horizontale. C'est vers ce dernier modèle que tendent les fournisseurs de services packagés, qui ambitionnent de prendre une place croissante dans la relation client.

L'importance de la **phase d'intégration** à l'environnement métier et au système d'information du client, et, le cas échéant, de localisation linguistique et culturelle, est régulièrement mise en avant. Un moteur d'intelligence artificielle ne peut, par exemple, pas être utilisé tel quel pour répondre aux besoins d'un client, et requiert un important travail en local chez l'utilisateur avant sa mise en service, car il doit pouvoir interagir avec un grand nombre de modules du système d'information et de données. Ce segment de la chaîne de valeur semble de ce fait mériter une attention toute particulière, puisqu'il conditionne la diffusion de ces outils, et représente aussi une opportunité de création d'activités et d'emplois plus difficilement délocalisables du fait du besoin de proximité et de maîtrise des spécificités nationales des équipes affectées à ces fonctions.

Alors que les progrès des moteurs de **recherche vocale** développés à l'étranger permettent désormais d'envisager, dans des conditions satisfaisantes, une utilisation à grande échelle dans divers contextes, les travaux menés en France, pourtant bien engagés, n'ont pas connu une issue aussi favorable. L'insuffisance des performances des solutions disponibles se manifeste, dans certains cas, par le recours à d'autres canaux de communication lors de la conception des robots (virtuels comme physiques). Il pourrait être pertinent, après analyse plus approfondie de l'acquis exploitable et des forces mobilisables, de réactiver et d'amplifier les chantiers recensés dans le cadre d'un programme ambitieux. Compte tenu du déport habituel du traitement des requêtes sur des serveurs distants, une telle réalisation présenterait un intérêt appréciable en termes de protection des données.

Il y a lieu de noter que la puissance publique est aussi une cible des applications de l'intelligence artificielle dans la relation à l'usager. L'intérêt est double : améliorer le service rendu et actionner le levier de la **commande publique** pour aider les entreprises à valider leur offre et se constituer un volant d'affaires pouvant faire office de socle pour leur projection à l'international.

Parmi les vecteurs de pénétration de l'intelligence artificielle dans les foyers, les **assistants domestiques vocaux** commencent leur percée (5 % aux États-Unis, arrivée en Allemagne). Les solutions les plus abouties sur ce créneau sont exclusivement étrangères, ce qui pose un vrai problème de souveraineté nationale. En principe, le moteur d'intelligence artificielle, auto-apprenant, est censé proposer les meilleures recommandations possibles à l'usager en privilégiant son intérêt, mais dans les faits, il est impossible de contrôler sa **neutralité** à l'égard des offres référencées, d'autant qu'à la différence d'un algorithme prédictif basé sur des statistiques dont les mécanismes peuvent être portés au jour, les motivations des choix d'un système faisant appel à l'intelligence artificielle seront quasiment impossibles à déchiffrer (en raison de leur plus grande complexité, de leur évolutivité au fur et à mesure de leur apprentissage, et de l'ampleur gigantesque de la masse de données traitées).
Si le combat relatif à la maîtrise des interfaces semble très compliqué sur le canal messagerie, il peut encore être gagné sur le créneau du **hub IA domestique**, qui pilote la vie connectée (domotique,

sécurité, consommation dont alimentation, santé…) en faisant éventuellement office de tiers de confiance, en raison de la faible pénétration de ce dispositif en France, et à condition d'une intégration réussie avec des domaines comme la santé ou bien les objets connectés.

Enjeux

La conservation du contrôle de la chaîne de valeur :
L'importance croissante qu'est appelée à prendre l'intelligence artificielle dans la relation client s'accompagne de l'implantation de **nouveaux entrants** dans la chaîne de valeur qui relie le producteur au consommateur, susceptibles de **s'emparer** des segments devenant les plus stratégiques en vertu de leur pouvoir d'orientation de l'acheteur vers telle ou telle offre (que le niveau de la rémunération, prélevée sur la marge du vendeur, accordée en contrepartie pourrait influer de manière prépondérante). Si, en particulier, une **convergence** des modes d'interaction avec les marques et les enseignes dans une toile de conversation adossée aux réseaux sociaux et aux messageries instantanées venait à se produire, le risque serait grand d'assister à un déplacement graduel de la **captation des données et de la valeur** vers les sociétés, majoritairement étrangères, exploitant ces plateformes, au détriment des acteurs économiques nationaux. Il en serait de même si les **assistants personnels et domestiques** venaient à se généraliser et prendre pied dans l'acte d'achat, et qu'aucune alternative aux services des géants du net (ex : Echo, Home) n'était proposée.
À défaut de pouvoir contrôler économiquement certains canaux de communication, plébiscités par la jeune génération, il importe de veiller à ce que les outils conversationnels, les bases de données qui leur font office de substrat et qu'ils alimentent, comme les processus métier pris en charge dans le cadre de l'automatisation des fonctions corporate, restent bien dans le giron des entreprises qui fabriquent et commercialisent les biens sous-jacents, ou des outsourceurs nationaux.

La protection et l'ouverture des données personnelles, et la transparence vis-à-vis de l'utilisateur :
Se pose la question de la **propriété**, de la **protection** et du **respect de la confidentialité** des données relatives aux consommateurs, auxquelles ont accès à la fois le gestionnaire de l'interface de communication et le marchand. Lorsqu'une transaction auprès d'un vendeur établi en France s'opère par l'intermédiaire d'une messagerie administrée depuis l'étranger plutôt que sur son site internet, le client n'est plus assuré de bénéficier de l'application des dispositions du droit communautaire et national. Un **cadre juridique**, tenant compte des bouleversements qu'induit la transformation numérique, doit être instauré, afin de sécuriser la collecte, le stockage et l'utilisation des données personnelles, tant pour le consommateur que pour les entreprises, tout en évitant de freiner le développement de l'intelligence artificielle, et de pénaliser les acteurs de taille modeste français par rapport aux poids lourds étrangers, qui sont plus à même d'emprunter des voies de contournement. Le durcissement des règles au sein de l'Union européenne en matière de protection de la vie privée ne doit pas provoquer de déséquilibre avec le reste du monde (États-Unis et Asie en particulier).
Outre les garanties juridiques, la mise en place de **tiers de confiance** pourrait apporter une réponse à ces préoccupations. Leur fonction est de s'intercaler entre la plateforme et le consommateur, en tant que gestionnaire d'identité, délégataire de la mise à disposition des données.
L'**ouverture** des données captées, par les plateformes en situation de quasi-monopole, auprès d'utilisateurs nationaux est un enjeu tout aussi important. Des entreprises françaises, dont des start-up, acquéraient ainsi la capacité de proposer des services basés sur ces données.
L'enjeu ne porte pas que sur la donnée, mais aussi sur la neutralité et l'explicabilité des **algorithmes**, comme évoqué ci-avant. L'utilisateur doit être informé qu'il a affaire à une intelligence artificielle.
Il en va de l'acceptation de ces solutions par le public, dont les enquêtes mettent en évidence les sentiments contrastés.

L'appropriation maîtrisée par le tissu économique
De même qu'Internet a provoqué une rupture dans la relation client, en permettant aux consommateurs de se renseigner plus facilement sur les produits, de comparer les prix et de consulter les commentaires des acheteurs, les évolutions de comportement et d'usage dont est vectrice, cette fois, l'intelligence artificielle, portent un risque élevé de **décrochage** des acteurs économiques qui n'auront pas su s'adapter à la nouvelle donne. Leur éviction des circuits commerciaux les plus fréquentés, et leur moindre capacité à procurer au client les services qu'il considèrera désormais comme allant de soi,

menaceront de façon de plus en plus préoccupante leur pérennité. La **sensibilisation** du tissu économique s'avère, dès lors, primordiale.
Un **accompagnement** du tissu économique est souhaitable pour l'aider à traverser avec succès cette mutation. La **maturité** des plus petites entreprises est souvent insuffisante, d'où leur propension excessive à calquer les solutions numériques sur les approches et méthodes traditionnelles sans révision profonde des processus.

L'anticipation de l'impact sur l'emploi et les compétences :
Les craintes suscitées par l'irruption de solutions logicielles et robotiques avancées à même de remplacer, au moins partiellement, des opérateurs humains, dans certains secteurs fortement pourvoyeurs d'emplois, doivent être surmontées au profit d'une réflexion plus générale sur l'incidence de ce phénomène sur l'**évolution** des processus métier, des organisations, des profils de poste et des compétences. Le consensus semble acquis sur le fait que les systèmes automatisés de traitement des interactions avec l'utilisateur requerront, au moins durablement, une **assistance humaine** pour pouvoir offrir une expérience satisfaisante (prise de relais de l'agent virtuel lorsqu'il atteint ses limites selon le principe d'escalade et actions concourant à son bon fonctionnement dont l'apprentissage). Les agents virtuels comme les robots physiques peuvent d'ailleurs être utilisés non pas en remplacement de conseillers mais en amont, en **appui** et en **complément**. Face à une mutation inéluctable, la stratégie gagnante consiste à accélérer la généralisation des systèmes intelligents et à miser sur les **nouveaux emplois** directs (conception, fabrication, mise en service et maintenance des solutions) et indirects (exploitation des outils et analyse et gestion des données) susceptibles d'être créés en France.
L'**anticipation** de l'évolution des qualifications recherchées doit intervenir au plus tôt pour maximiser les retombées sur le sol national. Une **communication** est à prévoir pour contrer les messages alarmistes sur les effets destructeurs sur l'emploi du développement de l'intelligence artificielle.
L'offre de **formation** doit être adaptée, pour proposer des profils de compétences très spécialisés mais aussi des profils plus transverses tournés vers le pilotage de la coopération des segments de la chaîne de valeur au service du client. L'objectif est de mettre en adéquation, quantitative et qualitative le vivier de compétences techniques et comportementales disponibles sur le marché du travail, en veillant à promouvoir la culture qui prévaut dans les entreprises innovantes.

Le développement de l'écosystème national de fournisseurs de technologies et solutions :
Il importe de parvenir à concilier l'**émergence** et le **développement** en France d'entreprises leader par leur **savoir-faire très pointu** sur une ou plusieurs briques, et la constitution d'acteurs capables d'un niveau élevé d'**intégration** pour pouvoir faire front aux plateformes étrangères. Il convient d'exploiter les **forces** françaises (notamment, le haut niveau des scientifiques et ingénieurs dans les disciplines impliquées, et la compétence en design et ergonomie des interfaces et parcours client), en évitant de se faire piller les talents, et de remédier à certaines faiblesses structurelles (financement notamment).
Les solutions actuelles sont très loin d'être parvenues à leur optimum. Des **développements technologiques** lourds sont encore requis avant que les systèmes commercialisés ne puissent être considérés comme étant arrivés à maturité quant à l'étendue des fonctionnalités proposées et au degré d'autonomie. Une intensification de l'effort de recherche et du transfert des résultats vers les entreprises est souhaitable.
Un enjeu fort réside dans la constitution de **pôles d'excellence** d'envergure. Il serait ainsi bénéfique d'encourager les synergies entre acteurs académiques et privés, dans l'optique de la mise au point de solutions abouties et commercialisables à partir de l'agglomération de leurs compétences respectives. Les grandes entreprises ont la capacité d'accompagner au plan managérial les plus petites structures et start-up pour qu'elles valorisent leur expertise, de partager avec elles leur intuition des attentes du client, de leur apporter du business, et de fédérer ces rapprochements.

L'éclosion d'un écosystème étendu :
Pour faire progresser les usages et l'offre en matière d'intelligence artificielle, les écosystèmes doivent associer les activités **connexes**.
À titre d'illustration, la compétitivité de la branche de la **robotique** consacrée à la relation client dépend de la performance et de la densité des relations entre les différentes parties prenantes : concepteur/fabricant/intégrateur de robots, fournisseurs de modules matériels (pour l'adjonction de capacités additionnelles) et logiciels (pour le fonctionnement autonome du robot et la gestion de ses

interactions), opérateurs de déploiement (adaptation, mise en service, administration), écosystème de financeurs et business angels, clients (en B to B et B to C).
L'intelligence artificielle a aussi un rôle à jouer dans la gestion des **objets connectés**, et a partie liée avec l'**internet des objets**, dont elle accroît la valeur ajoutée des éléments et des connexions.

Recommandations

Recommandation 1 : promouvoir, auprès du tissu économique, la composante relation client de la transformation numérique, en mettant en avant les approches fondées sur l'intelligence artificielle
L'accompagnement que les multiples opérateurs du développement économique exercent au quotidien auprès des petites et moyennes entreprises constitue une occasion privilégiée de faire acte, auprès de leurs responsables, de pédagogie et d'explication sur les opportunités et les risques que présente l'adaptation de leur politique client aux mutations à l'œuvre, d'autant que plusieurs études récentes pointent le retard pris par la France en la matière. Des manifestations et séances d'information ciblées pourraient également être organisées périodiquement, pour démystifier l'intelligence artificielle, lever les réticences infondées, et convaincre de la création de valeur à la clé.
Il serait utile d'élaborer, en complément de ces actions, un **livre blanc**. Conçu comme un guide méthodologique, il exposerait les scénarios d'utilisation et les fondamentaux d'un recours réussi à l'intelligence artificielle, ainsi que les précautions à prendre et les bonnes pratiques pour ne pas perdre le contrôle des données, savoir-faire et expertises métier stratégiques (d'une certaine manière confiés au système intelligent au cours de son apprentissage) tout en exploitant à leur plein potentiel les outils mis en place. Il comporterait aussi des témoignages et retours d'expériences. Face à la menace que représente l'attraction sur des carrefours d'audience administrés depuis l'étranger, comme les messageries instantanées, d'une part croissante des contacts entre les marques et les consommateurs, il préconiserait l'arme de l'originalité et de l'exclusivité de l'expérience client susceptible d'être proposée sur les autres canaux (site internet, application mobile) à l'opposé de la standardisation et de l'uniformité des interactions que ces interfaces unifiées imposent.
Cet ouvrage gagnerait à être prolongé par un **annuaire en ligne** des opérateurs à même d'accompagner les entreprises dans la modernisation de leurs processus et de leurs systèmes de gestion de la relation client (sociétés de conseil et de service informatique proposant éventuellement des prestations d'intégration, centres de contacts clients…). Ce référencement, catégorisé par nature d'intervention, s'enrichirait progressivement des avis postés par les bénéficiaires des prestations.
Des démarches collectives pourraient être suscitées et soutenues au sein de certaines professions dans le but de mutualiser une partie des travaux, technologies et données qu'implique l'usage de l'intelligence artificielle, dans les limites qu'impose leur caractère différenciant et donc stratégique sur le plan concurrentiel (comme, s'agissant des enseignes, l'enrichissement des descriptions de produits actuelles par des propriétés se rapportant directement aux usages et à l'expérience du client). Pourrait être encouragée la création de **plateformes d'achat sectorielles** donnant un accès unifié aux offres entrant dans un même univers de consommation. Elles pourraient être le fait d'une initiative commune des enseignes actives sur le segment considéré, ou bien de tiers se spécialisant sur ce nouveau service.

Recommandation 2 : faire valoir au plan européen et conduire à l'échelle nationale une adaptation pertinente, car conciliant défense des intérêts des consommateurs et compétitivité des entreprises nationales sur la scène mondiale, du cadre juridique et de l'appareil administratif de contrôle aux implications de la montée en puissance de l'intelligence artificielle dans la relation client
Il s'agit, à l'égard des **données**, de trouver un juste **équilibre** entre protection de la vie privée, faculté de personnalisation et d'enrichissement du service rendu grâce à une meilleure connaissance du client, élimination du risque juridique et donc financier pour les entreprises, et développement de l'offre et de l'utilisation en France de solutions faisant appel à l'intelligence artificielle. L'encadrement mis en place ne doit pas handicaper les entreprises françaises par rapport à leurs concurrents étrangers, et a contrario, leur ouvrir l'accès aux données que les grandes plateformes recueillent auprès des consommateurs nationaux (en faisant notamment jouer les règles nationales en matière de droit d'auteur). La création, au niveau communautaire ou national, d'une **instance de dialogue** réunissant les parties prenantes et la puissance publique, serait un moyen de se prémunir contre le risque d'adoption de restrictions excessivement pénalisantes pour le particulier ou pour l'écosystème de

fournisseurs de solutions, et d'infléchir les textes à venir dans le sens d'une indexation du phénomène à notre culture plutôt que d'une obstruction aussi inefficace que contre-productive.
Le cadre juridique doit aussi garantir la **neutralité**, la conformité aux dispositions visant à éviter les pratiques anti-concurrentielles, et l'**explicabilité** et la traçabilité des choix des **algorithmes** sous-jacents. Il pourrait être envisagé également d'imposer que l'utilisateur soit systématiquement informé qu'il a affaire à une intelligence artificielle.
Recommandation 3 : encourager le recours à l'intelligence artificielle par la puissance publique dans ses rapports à l'usager
L'objectif est de positionner l'État (et les collectivités) non seulement en régulateur mais aussi en amorceur et en consommateur de solutions faisant appel à l'intelligence artificielle, dans le cadre de la modernisation de l'action publique, et par la faculté du service public à faire office de terrain d'expérimentation et de validation de solutions innovantes. À l'instar d'autres pays, la France a tout à gagner à mobiliser la **commande publique** en tant qu'instrument efficace pour soutenir l'émergence et le développement d'entreprises leaders sur leur marché.
Une **sensibilisation** des décideurs, des personnels en poste et des nouveaux entrants, à partir d'expériences concluantes, un renforcement de leur connaissance des procédures appropriées aux commandes impliquant une part significative de recherche et développement instaurées ces dernières années, ou encore le lancement de marchés privilégiant des réponses novatrices basées sur des briques d'intelligence artificielle conçues en France sont à considérer.

Recommandation 4 : valoriser les opportunités de carrière dont est porteuse l'application de l'intelligence artificielle à la relation client, en œuvrant en faveur d'une compréhension partagée et apaisée des mutations des emplois à venir, et réaménager en conséquence l'offre de formation initiale et continue, ainsi que le droit social
Une **communication** est à prévoir pour contrer les messages alarmistes sur les effets destructeurs pour l'emploi de la propagation des usages de l'intelligence artificielle (sans nier que le pas supplémentaire que ces technologies sont en mesure de faire franchir à la poursuite de l'automatisation des tâches routinières est susceptible de se traduire localement par des suppressions de postes à dominante opérationnelle). Est à mettre en exergue l'argument du gain de performance que retirent les entreprises de l'adoption de ces solutions, et donc du bénéfice que ces dernières apportent à l'ensemble de l'économie en termes de capacité de résistance face à la concurrence et de rythme de croissance. Elle doit être adossée à une **campagne d'information positive** sur les besoins en personnel résultant des bouleversements survenant au sein des processus constitutifs de la gestion de la relation client et de l'émergence de nouvelles activités économiques. Il convient de caractériser ces **métiers d'avenir**, en termes notamment de qualifications requises et de possibilités de reconversion des effectifs en place (simple utilisateur de solution, expert métier entraîneur ou superviseur, responsable de projet de déploiement…), et de les faire connaître auprès des jeunes comme des salariés dont les fonctions sont menacées.
Une **rénovation des formations** initiales et continues afférentes aux emplois affectés est indispensable, les contenus devant incorporer au plus vite les impacts de l'intelligence artificielle sur les modalités d'exercice des fonctions concernées, et préparer les opérateurs à la coopération avec les systèmes intelligents. La mise en place de cursus, à dominante scientifique, économique ou littéraire, ou bien hybrides, spécifiques aux nouveaux profils recherchés, est aussi à prévoir.
Après une évaluation transverse plus approfondie des recrutements attendus au cours des prochaines années, il y aurait également lieu d'accroître le flux sortant d'ingénieurs et techniciens spécialisés, en favorisant les enseignements couvrant plusieurs briques technologiques et leur croisement, et procurant des compétences multiples (comme développeur et data scientist par exemple).
Les questions que soulève le déplacement du centre de gravité de la collaboration humain/machine et la subordination des travailleurs aux plateformes nécessitent d'être résolues dans un cadre plus général que celui de la seule relation client.

Recommandation 5 : stimuler l'écosystème français élargi de concepteurs, fournisseurs et intégrateurs de modules, solutions et services déclinant l'intelligence artificielle dans le champ de la relation client, en favorisant les synergies
Les politiques publiques menées depuis plus d'une décennie pour favoriser les rapprochements entre entreprises, laboratoires publics, organismes d'enseignement supérieur et technique, financeurs et

investisseurs providentiels, et clients prêts à accueillir des expérimentations doivent, de manière générale, être orientées en direction des disciplines et usages que recouvre l'intelligence artificielle. Il en va de même des dispositifs en faveur de l'amorçage, de la création et du développement des start-up. Le lancement d'appels à projet sur des segments considérés comme porteurs est, en particulier, un moyen efficace de consolider des partenariats autour de chefs de file disposant de moyens humains, techniques et financiers à due proportion des ambitions poursuivies. Les actions et instruments mis en œuvre devront inclure les secteurs connexes, tels que la robotique et les objets connectés.
Il conviendrait de mieux faire connaître les innovations françaises, tout particulièrement les produits commercialisés par des petites et moyennes entreprises, en mettant en avant, outre leurs atouts propres, les avantages qu'elles présentent en termes de souveraineté nationale. La création d'un **site portail** référençant, thématique par thématique, les **solutions** existantes, et facilitant le choix de l'offre la plus adaptée à chaque situation, au moyen de tableaux comparatifs de leurs fonctionnalités et de leurs modalités de déploiement, semble être une approche judicieuse. L'objectif passe aussi par une étroite coopération avec les intégrateurs, qui, situés en aval de la chaîne de compétences, effectuent un travail fondamental de design de solutions sur mesure (en qualité de professionnels de la relation homme-machine), et interviennent chez le client pour la mise en service, en procédant aux connexions avec les nombreux modules et données impliqués du fait du caractère pervasif de l'intelligence artificielle.

Recommandation 6 : faciliter l'accès aux technologies développées en France sous-jacentes aux applications de l'intelligence artificielle à la relation client

De manière à maximiser les retombées sur le sol français, réduire la dépendance vis-à-vis des grands groupes étrangers, et étoffer et démocratiser l'offre de solutions, il est essentiel d'améliorer, outre leur performance intrinsèque et leur interopérabilité, l'accessibilité des briques technologiques nationales mobilisables pour la conception d'outils ou de services. Il apparaît opportun, à cet effet, d'envisager, en parallèle des actions concourant à faire progresser la maturité des technologies disponibles, la création d'une **banque d'algorithmes et de bibliothèques de codes** mis à disposition des développeurs (sociétés spécialisées ou bien équipes dédiées au sein des marques). Le modèle économique et les conditions d'utilisation restent à déterminer, l'essentiel étant que le système demeure ouvert, sans leadership d'acteur dominant. L'objectif est de proposer une panoplie de services cognitifs, accessibles sous forme de cadriciel ou via des interfaces de programmation applicative, en fournissant éventuellement un environnement de test.
Sont à privilégier les approches à même d'expliciter les motivations des choix faits par le système (éviter l'effet boîte noire et l'effet bulle de filtres), prévoyant la portabilité dans les objets, et la maîtrise de la localisation du stockage des données et du traitement des requêtes, lorsque ces opérations sont déportées sur des serveurs distants. La France aurait qualité à prôner et devenir chef de file de la mise en place, à l'échelle européenne, d'un système de **labellisation** garantissant le respect d'une charte énonçant les principes jugés inhérents au bon usage de l'intelligence artificielle.

Recommandation 7 : se positionner sur l'un des vecteurs majeurs de pénétration de l'intelligence artificielle dans le quotidien du consommateur, les assistants de vie

La finalité poursuivie est de faire émerger, à l'intérieur de nos frontières ou au travers d'un partenariat européen, un **référent du hub IA** des foyers français, qui pilote leurs interactions avec les services et objets connectés (domotique, sécurité, consommation dont alimentation, santé…) et donc une part, appelée à croître au fil du temps, de leurs relations avec les marques, en faisant éventuellement office de tiers de confiance. L'**assistant** embarqué n'a pas seulement vocation à être installé à domicile, mais trouverait aussi sa place sur des bornes implantées dans les lieux accueillant du public et constituant un point d'accroche pour la relation client, ou encore sur les « wearables ».
Les équipementiers électroniques et opérateurs des postes et télécommunications (fournisseurs d'accès internet) pourraient être parties prenantes d'un tel projet. Un **label**, destiné à susciter la confiance chez le consommateur, pourrait être créé en direction de l'écosystème de start-up associées à la station. Il marquerait la distance prise par rapport aux dispositifs étrangers sur le plan éthique, du point de vue de l'utilisation des données personnelles recueillies, tout en permettant un modèle économique viable supposant la valorisation de ces informations.
Au travers d'une deuxième génération d'assistant, l'interaction entre le consommateur et les agents conversationnels des marques pourrait tendre vers un dialogue entre intelligences artificielles au

service respectif du vendeur et du client, chaque système maîtrisant les attentes de la partie dont il représente les intérêts et se chargeant de sélectionner et de négocier pour son compte.
Deux **appels à projet** pourraient être lancés dans le cadre du programme d'investissement d'avenir, au titre des volets correspondant à leur TRL respectif, pour susciter la constitution de consortiums.

Annexe 1

Identité des contributeurs

Pilotes du groupe de travail :
- Éric Dadian (Président de l'Association française de la relation client)
- Cyril Madar (Chef de bureau des services pour la performance compétitive à la Direction générale des entreprises, Ministère de l'économie et des finances)

Participants au groupe de travail :
- Association des agences-conseil en communication : Catherine Michaud (Administratrice, Présidente de la délégation Customer Marketing)
- Société Do You Dream Up : Souad Zarat (Responsable des ventes et des partenariats), Mathieu Changeat (Gérant)
- Société Calldesk : Aimé Dushimire (Co-fondateur)
- Société Antvoice : Alban Peltier (Directeur général), Rodolphe Mirilovic (Directeur technique), Simon Lefebvre (Responsable data science)
- Société Voyages-SNCF : Michaël Pépin (Responsable performance et projets), Thibault Meunier (ex-Responsable des partenariats stratégiques)
- Société AccorHotels : Fabrice Otano (Directeur des données), Philippe Périé (Vice-président chef de l'analyse)
- Société Hease Robotics : Jade Le Maître (Directrice technique et co-fondatrice)
- Société Proxem : François-Régis Chaumartin (Président)
- Pôle de compétitivité des industries du commerce : Pierre Blanc (Responsable innovation grands comptes)
- Direction générale des entreprises, Ministère de l'économie et des finances : Olivier Grollet (Chargé de mission relation client par intérim), Dominique Patebex (Chef du bureau des professions du tourisme), Brigitte Sica-Barcella (Chargée de mission marketing publicité)

Personnes extérieures au groupe de travail consultées :
- Société IBM France : Jean-Philippe Desbiolles (Vice-président Watson)
- Société Cap Gemini : Jean-Claude Guyard (Directeur « lab applied innovation exchange » France)
- Société iAdvize : Nicolas de Rosen (Directeur projet stratégique)
- Société Leroy Merlin : Albin Dorian-Juliard (Directeur ventes et relation client)
- Société Voyage Privé : Laurent Dupé (Vice-président marketing international)
- Société Allo-Media : Frédéric Daniel (Directeur général adjoint)
- Société Webhelp : Matthieu Bouin (Directeur de la stratégie)

Documents

- définition de la relation et du parcours client, et impact des technologies numériques et d'intelligence artificielle (association française de la relation client et pôle de compétitivité des industries du commerce)
- les technologies mises en jeu par les applications de l'intelligence artificielle à la gestion de la relation client
- les nouveaux métiers liés à la transformation numérique de la relation client et à son enrichissement par l'intelligence artificielle et la robotique
- récapitulatif des actions à portée opérationnelle figurant dans les recommandations

Annexe 2 - définition de la relation et du parcours client, et impact des technologies numériques et d'intelligence artificielle

La Relation Client regroupe l'ensemble des dispositifs ou opérations de marketing et de support ayant pour but d'optimiser la qualité de l'expérience client, de fidéliser et de maximiser le chiffre d'affaires ou la marge par client.

La gestion de la relation client ou GRC regroupe à la fois des techniques d'analyse des données clients, des opérations marketing et des opérations de support produits.

La relation client représente donc l'ensemble des interactions qui utilisent tous les canaux de contacts disponibles avec le client, comme le face à face dans les boutiques mais aussi le téléphone, le mail, le chat, la vidéo, whatsapp, messenger de facebook, twitter à distance.

Depuis l'époque où l'achat d'un produit se faisait uniquement en boutique et avec l'aide d'un vendeur, les modes de consommation et de distribution ont bien changé. De nouvelles technologies, des concepts commerciaux et marketing ont fait leur apparition révolutionnant les habitudes des consommateurs et leurs attentes. Plus question d'envisager un achat sans passer par internet, les réseaux sociaux ou un téléconseiller par téléphone. Le client d'aujourd'hui est « connecté » et multicanal: il n'hésite pas à solliciter l'entreprise via plusieurs « canaux de communication » pour trouver entière satisfaction.

Avec les années 2000, l'internet révolutionne l'économie et les usages. On assiste à l'émergence d'un 3e canal d'interaction majeur. Une vraie rupture s'opère dans la relation client. Les consommateurs ont désormais l'opportunité de se renseigner plus facilement sur les produits, de comparer les prix avant d'entrer en contact avec les marques de leur choix. Parallèlement à l'internet, le mobile s'émancipe, notamment au travers d'usages nouveaux tels que le SMS. Les entreprises renforcent leurs stratégies de segmentation et de fidélisation grâce au marketing dit « pull ».

En très peu d'années, les consommateurs se sont équipés de smartphones et de tablettes connectées à Internet, des « applications » sont apparues et n'ont cessé de se développer ; nos modes de vie en sont bouleversés car au lieu de nous déplacer pour acheter des biens ou bénéficier des services dans des lieux dédiés, de nouvelles chaînes de production à base d'échange d'informations organisent la mise à disposition de savoirs et de biens dont nous avons besoin. Il s'agit d'une **véritable révolution client**, d'une autre façon de satisfaire nos besoins, « *user centric* ». L'« uberisation », si vite entrée dans le langage courant, en est le meilleur exemple.

Le numérique change la perception du travail, notre rapport au temps, les notions de distances et donc les manières de travailler ! Tout ce qui peut être virtualisé dans la vie de tous les jours, le devient, et ces représentations « digitales » du monde réel sont partagées, reproduites, transformées par des consommateurs. Les robots, les objets connectés, la réalité virtuelle, l'IA, nous font passer d'un consumérisme passif à **une fabrique collective centrée sur le client et le faire soi-même.**

Nous changeons vers une société « démultipliée, augmentée » où tout devient **relationnel**, et où le client « devenu expert », veut échanger, partager de la connaissance autour d'un produit, d'une marque et de ses valeurs avec sa communauté avant de se décider d'acheter.

Le commerce a été le premier secteur concerné. Des entreprises, comme Amazon dès 1994, ont proposé des applications qui permettent d'acheter toutes sortes de biens sans aller dans des boutiques et en ayant un choix sans commune mesure par rapport à celui que l'on a en y allant. Le e-commerce s'est mis à menacer le commerce traditionnel.

Ce phénomène a ensuite été élargi aux services, sous des formes diversifiées. Pour les services d'hôtellerie, Booking.com par exemple proposent dès 2003 des « solutions » permettant d'accéder à l'ensemble de l'offre de chambres disponibles dans tous les hôtels. En 2012, Airbnb propose que ce soient les habitants eux-mêmes qui puissent mettre à disposition temporaire leurs chambres ou même leurs logements. Des sociétés, comme Uber en 2009, mais avant elle Blablacar ou Leboncoin, en 2006

ont proposé d'apparier l'offre et la demande de biens entre consommateurs (voitures, produits d'occasion…) en faisant prendre en charge par les consommateurs eux-mêmes leur mise à disposition.

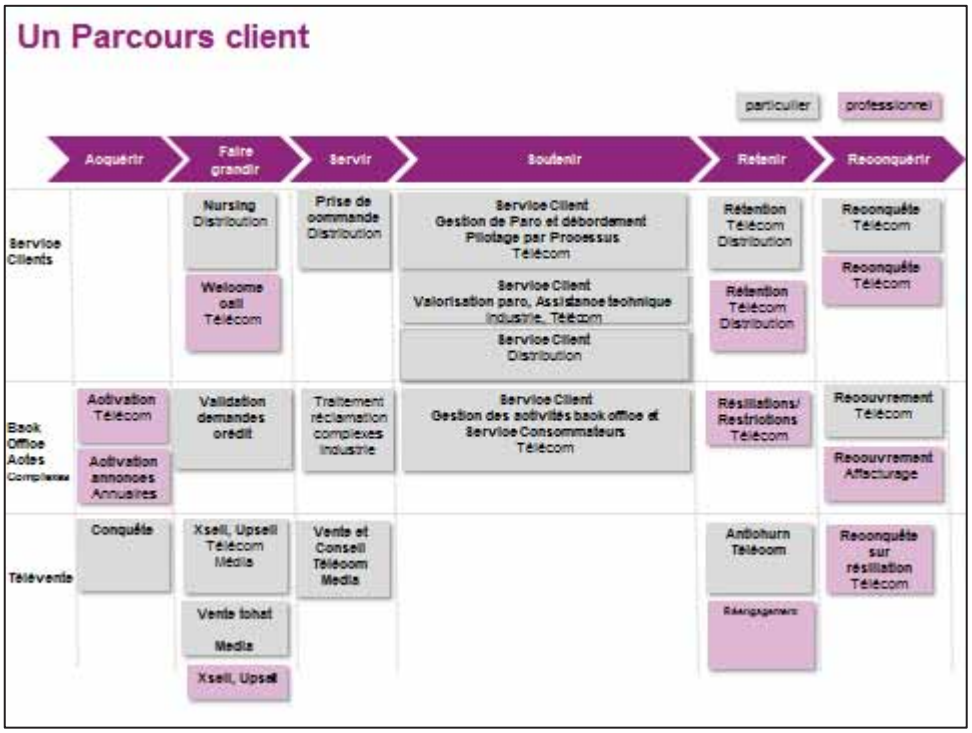
Avec le développement des objets « connectés », on peut tout savoir de la vie des consommateurs à distance et en continu : capteurs de fumée, de température, de pas effectués, de pouls, d'humidité des plantes, d'ouverture des portes et fenêtres, du chauffage… L'humain devra, plus que jamais, faire preuve de discernement et d'initiative face à ces nouveaux flux d'information.

L'enjeu pour les marques sera de construire une **relation client attentionnée** entre les marques et leurs clients, et ce, malgré la distance permise par Internet, l'empathie, la bienveillance, la gentillesse devient une valeur première pour réassurer, fidéliser, satisfaire un client, apporter l'émotion qui fidélisera le client.

Le parcours client :

Si la stratégie multicanal consiste à prendre en compte chaque point d'entrée, web, centre de contacts, réseaux sociaux, boutique physique, le parcours client invite, quant à lui, à une réflexion plus globale sur les outils à mettre en place pendant les trois phases de relation avec l'entreprise :

- la recherche d'informations avant l'achat d'un produit ou d'un service,
- l'acte d'achat lui-même
- le suivi après l'achat.



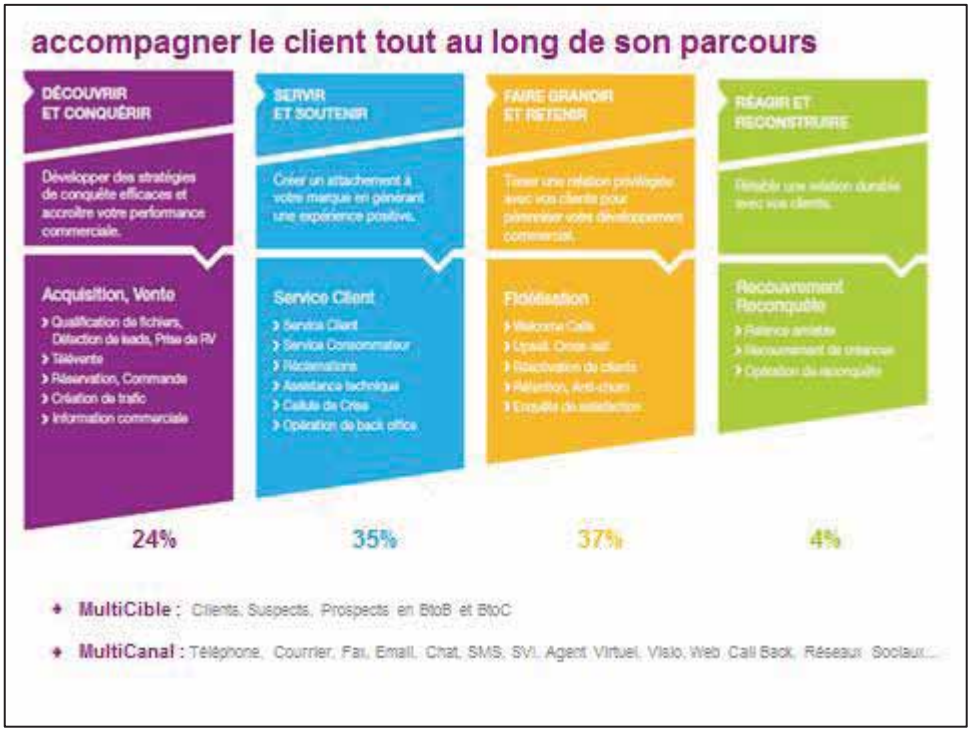
On pourrait illustrer un parcours typique ainsi : le client s'informe en premier lieu sur internet, puis se déplace en boutique, avant de revenir comparer les prix sur internet, et lorsqu'il se décide à acheter, il appelle le call center pour plus d'informations.

La modélisation de ce parcours vise la recherche systématique d'une parfaite fluidité entre les canaux et à l'intérieur d'un même canal. Le but : proposer une relation client sans rupture et gagner la satisfaction des clients.

Une étude AFRC Médiamétrie récente (2016) montre à quel point Internet est devenu le média incontournable pour démarrer son parcours client, puisque parmi les internautes français, 81 %

préparent leurs achats sur la toile. Sur un plan statistique, cette étude a montré qu'en moyenne, en 2015, un client a eu recours à un service client via 3,6 canaux différents (vs 2,6 en 2009). Les principaux canaux d'un parcours client restent : Internet (69 % vs 59 % en 2015), le téléphone pour la première prise de contact (7 % vs 18 % en 2015), le face à face/boutique (10 % vs 18 % en 2015) et l'entourage (8 % vs 12 % en 2015).

Dans tous les cas, la marque incarnée par le ou la conseiller(e), se met à MA disposition, moi client. La **posture de réassurance**, d'**écoute** sont des exemples d'empathie.



Le concept de « courbe d'émotion » part d'un constat simple : plus l'émotion pour le client est forte, plus l'entreprise doit le rassurer. Plus l'émotion est faible, plus l'entreprise doit l'attirer vers elle. Dans l'idéal, il faut donc proposer des services adaptés à chaque étape de la courbe d'émotion pour créer un lien entre les émotions du client et l'entreprise à chaque étape du parcours client.

Les apports de l'IA dans la relation client :

IA comme moyen de répondre à de nouvelles pratiques du commerce

Le parcours d'achat est cross-canal avec une imbrication physique / digital, d'où une multiplication des données sur les achats (CRM), les comportements (tracking), les conversations (réseaux sociaux), les émotions (à venir) => nécessité de raisonner globalement en quasi temps réel avec des grandes quantités de data issues de différentes sources et redistribuées sur tous les canaux.

L'IA se substitue aux approches antérieures de BI / data mining plutôt utilisées en back-office.

IA comme moyen de répondre à de nouvelles exigences du commerce : immédiateté et personnalisation de la relation avec le consommateur

L'IA permet de :

- 1. développer le marketing prédictif pour gérer le paradoxe d'une « personnalisation de masse de la relation » (approche push)
- 2. fournir au consommateur les moyens d'exprimer ses besoins (éventuellement latents) à un commerçant qui doit structurer une réponse commerciale adaptée (approche pull)

Besoin		Potentiel de l'IA
Comprendre l'intérêt, l'envie, l'intention, le projet mais aussi l'émotion (car il est nécessaire de positionner le débat à un autre niveau que la seule mise en avant du produit)	Comprendre le comportement d'achat	- analyse des pratiques et habitudes d'achat du consommateur - caractérisation du consommateur en situation d'achat : âge, genre - détermination des patterns de comportement d'achat
	Comprendre la conversation au téléphone, sur site et sur les réseaux sociaux Comprendre la requête e-commerce exprimée avec les mots du consommateur	- reconnaissance vocale - analyse sémantique - traitement du langage naturel
	Comprendre le contexte du consommateur	- reconnaissance d'objets et de situations
	Comprendre l'émotion, la personnalité	- analyse des expressions faciales - analyse des paramètres physiologiques - analyse de la voix - analyse de l'expression écrite
Déterminer le bon produit / service aux bons moment, lieu, étape du parcours d'achat	(info à destination du client ou du vendeur)	- ontologies de produits (méta-données associées aux produits) - moteur de recommandation
Interagir / dialoguer automatiquement et sur le bon média d'interaction		- synthèse vocale - génération auto de texte - restitution des émotions - chatbot
Anticiper les besoins	Comprendre le comportement de consommation	- analyse des pratiques et habitudes de consommation - caractérisation du consommateur en situation de consommation : âge, genre, famille - détermination des patterns de comportement de consommation
Proposer un majordome électronique de la vie quotidienne	(type Alexa, Google Now, Cortana, Siri, Samsung Viv, …)	- agent conversationnel

Annexe 3 - les technologies mises en jeu par les applications de l'intelligence artificielle à la gestion de la relation client (à titre indicatif)

	Agent conversationnel vocal	Agent conversationnel textuel	Robot d'interaction	Recommandation prédictive	Analyse textuelle	AI monitoring/testing/survey
Reconnaissance vocale	X		X			
Reconnaissance visuelle			X			X
Reconnaissance olfactive/gustative/tactile			X			X
Reconnaissance biométrique	X		X			X
Psychologie cognitive	X	X	X	X	X	X
Modélisation bayésienne / réseaux de neurones	X	X	X	X	X	X
Calcul de similarité	X	X	X	X	X	X
Traitement du langage naturel	X	X	X		X	X
Moteur de dialogue	X	X	X		X	
Analyse sémantique					X	X
Analyse de données hétérogènes			X	X		X
Analyse prescriptive	X	X	X	X		
Détection d'intentions et émotions	X	X	X	X		X
Décision sous contraintes et incertitude			X	X		X
Recherche sur données massives				X	X	
Planification			X	X		X
Robotique et déplacement			X			
Collaboration humain/machine			X			
Interface machine/machine, protocoles	X	X	X	X	X	X
Apprentissage automatique	X	X	X	X	X	X
Créativité artificielle	X	X	X			
Auto-programmation			X			X
Explicabilité de l'apprentissage	X	X	X	X		
Explicabilité des raisonnements				X	X	X
Présentation des résultats				X	X	X
Déontique	X	X	X	X		
Sécurité des données et des échanges	X	X	X	X	X	X

Annexe 4 - les nouveaux métiers liés à la transformation numérique de la relation client et à son enrichissement par l'intelligence artificielle et la robotique (liste non exhaustive)

Directeur de la transformation digitale (chief digital officer) : responsable, placé de préférence au sein de l'équipe de direction, de la mutation numérique au sein d'une organisation (évaluation de la maturité, orchestration du renouvellement de la stratégie, de la relation client et des processus internes, conduite du changement, accompagnement des équipes, maintien à niveau une fois la transition initiale achevée…)

Chef de projet digital : responsable de la conduite de projets de transformation numérique, dont il orchestre et supervise, dans le respect des objectifs de coût, de délai et de fonctionnalités, la préparation et la mise en œuvre

Ingénieur en intelligence artificielle : chargé de la conception et de la supervision du développement puis du déploiement de systèmes intelligents dans un contexte donné, capables de se substituer ou de venir en appui aux opérateurs, en s'assurant de la bonne coopération homme machine

Développeur ou analyste programmeur : chargé de la conception des algorithmes et de l'écriture ou de l'adaptation du code informatique en vue de la réalisation de logiciels, d'applications, d'interfaces…

Analyste usage (user analyst / user scientist) : chargé de la caractérisation des besoins et des attentes du consommateur en vue de la rédaction du cahier des charges fonctionnel de la solution à base d'intelligence artificielle

Analyste données (data analyst / data scientist) : chargé du choix, de l'architecture et de l'exploitation des bases de données en vue de répondre à une problématique déterminée ou de dégager des enseignements inattendus, grâce à l'emploi de techniques de modélisation statistique, de segmentation, d'établissement de corrélations et de restitution sous forme d'aide à la décision

Analyste de corpus textuel (text scientist) : chargé de superviser les ressources linguistiques (lexicales, syntaxiques, sémantiques et pragmatiques) d'un système d'analyse textuelle, obtenue notamment par apprentissage automatique sur corpus, il a une double compétence en étant simultanément infolinguiste et data scientist

Délégué à la protection des données (data protection officer) : successeur du correspondant informatique et liberté, obligatoire à partir de fin mai 2018 dans le secteur public et dans toute organisation procédant à des traitements à grande échelle de données personnelles, chargé du respect du droit et de la vérification de la conformité des systèmes et procédures d'exploitation

Designer d'expérience utilisateur (UX designer) : chargé de l'optimisation de l'expérience utilisateur sur le plan de l'ergonomie des interfaces numériques

Responsable commerce électronique : chargé de la performance du canal ventes en ligne en relation avec les autres canaux de vente, au travers de la gestion du référencement, de l'utilisation des leviers du marketing, de la publicité et de la présence sur les réseaux sociaux, de l'amélioration des fonctionnalités de la plate-forme, de l'augmentation du taux de conversion et de la fidélisation des visiteurs

Gestionnaire de communauté et de réputation (community et reputation manager) : porte-parole d'une organisation ou d'une marque sur le web et les réseaux sociaux, chargé d'en accroître la notoriété, d'en soigner la réputation, d'interagir avec les visiteurs, d'animer et modérer les conversations, de créer du lien avec les consommateurs ou usagers, de diffuser du contenu informationnel ou promotionnel, d'effectuer une veille sur les avis et commentaires, de lancer des événements…

Responsable du trafic (traffic manager) : chargé de l'acquisition de trafic vers le site par la maximisation de sa visibilité et le choix des lieux de référencement et de positionnement de liens publicitaires les plus pertinents

Référent apprentissage d'intelligence artificielle : chargé de l'éducation/entraînement initial et en exploitation du système intelligent, de la correction en temps réel ou a posteriori de ses erreurs, du contrôle de son éducation

Responsable de la sécurité des systèmes d'information : chargé de garantir la sécurité physique et logique, la disponibilité et l'intégrité des infrastructures, des applications et des données, en définissant la politique de l'organisation en la matière, après avoir analysé les risques encourus (en tenant compte des implications du recours à l'intelligence artificielle et à la robotique), en veillant à son application, en menant des actions de sensibilisation en direction des utilisateurs, et en identifiant et en corrigeant en continu les vulnérabilités

Ingénieur cogniticien : chargé de la prise en compte de la dimension interface homme/machine lors de la conception de systèmes ou de sites, en mobilisant sciences humaines et techniques pour adapter les fonctionnalités aux attentes, optimiser la coopération, perfectionner l'ergonomie et augmenter la sécurité et le confort, en réalisant au besoin des tests auprès d'utilisateurs

Sculpteur d'intelligence artificielle : chargé de conceptualiser la personnalité de l'intelligence artificielle (valeurs, attitudes, comportements…)

A/B testeur en intelligence artificielle : chargé du contrôle de version (versioning) et de l'évaluation des modèles d'intelligence artificielle avant ou pendant la commercialisation

Directeur de la robotisation (chief robotics officer) : responsable de la conception et de la mise en œuvre de la stratégie d'automatisation et de robotisation de processus au sein d'une organisation (détection des opportunités, sélection des solutions techniques pertinentes, évaluation des travaux d'intégration nécessaires, anticipation des résistances internes et accompagnement des équipes au changement, construction, soumission et conduite du programme dans le respect de la réglementation, supervision de l'exécution, surveillance en fonctionnement, évolutions …)

Designer d'interaction homme-robot (RX designer) : chargé de la définition des modalités de la relation entre le robot et les utilisateurs en fonction des services attendus, des publics ciblés, des enjeux de sécurité et du contexte d'exploitation, puis de la conception, du test et de la validation des différentes interfaces

Superviseur cycle de vie d'intelligence artificielle : chargé de la vérification de la qualité de service globale d'une technologie délivrée au marché

Gestionnaire de parc de robots : chargé de l'administration de la cohorte de robots d'interaction en service dans un lieu accueillant du public (acquisition, maintenance, évolution…)

Technico-commercial en intelligence artificielle : chargé de la gestion de la mise sur le marché, de la vente et du suivi des solutions recourant à l'intelligence artificielle

Conseiller en robot : chargé de caractériser et sélectionner le robot répondant à l'usage et à l'environnement auxquels il est destiné

Juriste spécialisé en intelligence artificielle : chargé d'élaborer, évaluer et/ou appliquer le corpus juridique relatif à l'intelligence artificielle et au droit des robots

Annexe 5 - récapitulatif des actions à portée opérationnelle figurant dans les recommandations

Intitulé	Contenu
Élaboration d'un livre blanc sur le recours à l'intelligence artificielle pour la gestion de la relation client, accompagné d'un annuaire en ligne des prestataires pouvant assister les entreprises dans cette démarche	Ce guide méthodologique exposerait les scénarios d'utilisation et les clés de la réussite, ainsi que les précautions à prendre et les bonnes pratiques pour ne pas perdre le contrôle des données, savoir-faire et expertises métier stratégiques. Il comporterait des retours d'expériences. Face à la menace de l'attraction sur des carrefours d'audience administrés depuis l'étranger d'une part croissante des contacts entre marques et consommateurs, il préconiserait l'arme de l'originalité de l'expérience client susceptible d'être proposée sur les canaux traditionnels à l'opposé de l'uniformité des interactions que ces interfaces unifiées imposent. Cet ouvrage gagnerait à être prolongé par un annuaire en ligne des opérateurs à même d'accompagner les entreprises dans la modernisation de leurs processus et systèmes de gestion de la relation client. Ce référencement, catégorisé par nature d'intervention, s'enrichirait progressivement des avis postés par les bénéficiaires des prestations.
Incitation à la création de plateformes d'achat sectorielles intelligentes françaises	Pourrait être encouragée, afin de contrer les visées des gestionnaires de messagerie instantanée, la création de plateformes nationales d'achat sectorielles, tirant parti de l'intelligence artificielle pour la gestion des interactions avec le consommateur (expression du besoin en langage naturel, recommandations, assistant shopping), donnant un accès unifié aux offres entrant dans un même univers de consommation, au moyen d'une interface façonnée sur mesure. Elles pourraient être le fait d'une initiative commune des enseignes actives sur le segment considéré, ou bien de tiers se spécialisant sur ce nouveau service. Elles donneraient lieu à la mutualisation d'une partie des travaux, technologies et données, comme l'enrichissement des descriptions de produits par des propriétés se rapportant directement aux usages et à l'expérience du client, dans les limites qu'impose leur caractère différenciant et donc stratégique sur le plan concurrentiel.
Soutien à la création d'un site portail référençant les solutions françaises en matière d'intelligence artificielle appliquée à la relation client	Dans le but de mieux faire connaître les innovations françaises, tout particulièrement les produits commercialisés par des petites et moyennes entreprises (qui ne disposent pas de budgets marketing et de forces de vente équivalents à ceux des grands groupes étrangers), il serait utile de mettre en place un site portail référençant, thématique par thématique, les solutions existantes, et facilitant le choix de l'offre la plus adaptée à chaque situation, au moyen de tableaux comparatifs de leurs fonctionnalités et de leurs modalités de déploiement (ces distinctions sont souvent subtiles et difficiles à saisir pour un client potentiel peu familier avec les technologies). Ce site mettrait en avant les avantages qu'elles présentent en termes de souveraineté nationale. Il pourrait être géré par la communauté des développeurs, en étroite coopération avec les intégrateurs.
Soutien à la création d'une banque d'algorithmes et de bibliothèques de codes faisant appel à l'intelligence artificielle et exploitables pour la gestion de la relation client, couplée à un label éthique	De manière à améliorer l'accessibilité des briques technologiques nationales d'intelligence artificielle mobilisables pour la conception d'outils ou de prestations au service de la relation client, il apparaît opportun d'envisager, en parallèle des actions concourant à faire progresser la maturité et l'interopérabilité des technologies, la création d'une banque d'algorithmes et de codes de services cognitifs mis à disposition des développeurs (sociétés spécialisées ou bien équipes dédiées au sein des marques), en fournissant éventuellement un environnement de test. Le modèle économique et les conditions d'utilisation restent à déterminer, l'essentiel étant que le système demeure ouvert, sans leadership d'acteur dominant. La France aurait qualité à prôner et devenir chef de file de la mise en place, à l'échelle européenne, d'un système de labellisation garantissant le respect d'une charte énonçant les principes jugés inhérents au bon usage de l'intelligence artificielle.
Lancement d'appels à projet en faveur de l'émergence d'assistants intelligents de vie connectée, français ou européen, adossés à un écosystème de services labellisés	L'objectif est de faire émerger, à l'intérieur de nos frontières ou via un partenariat européen, un référent du hub IA des foyers français, qui pilote leurs interactions avec les services et objets connectés, et donc une part croissante de leurs relations avec les marques. L'assistant embarqué trouverait aussi sa place sur des bornes implantées dans les lieux constituant un point d'accroche pour la relation client ou sur les « wearables ». Les équipementiers électroniques et opérateurs des postes et télécommunications pourraient être parties prenantes d'un tel projet. Un label, destiné à susciter la confiance chez le consommateur, pourrait être créé en direction de l'écosystème de start-up associées à la station. Au travers d'une deuxième génération d'assistant, l'interaction entre le consommateur et les agents conversationnels des marques pourrait tendre vers un dialogue entre intelligences artificielles au service respectif du vendeur et du client.

GROUPE DE TRAVAIL 2.5
FINANCE

Groupe Finance

Etat des lieux

I Périmètre

Les investissements dans les start-up en intelligence artificielle (IA) appliquée à la finance ont dépassé 2 milliards de dollars en 2016, d'après PWC/Pitchbook. Ces start-up ressortent principalement de trois domaines technologiques :

- les logiciels auto-apprenants (apprentissage profond, réseaux neuronaux, machine learning…) ;
- le traitement automatique du langage naturel (NLP) ;
- les systèmes experts.

Cette taxonomie n'est bien sûr pas absolue : le *machine learning* est utilisé en NLP qui peut lui-même être utilisé pour faire des systèmes experts. Mais elle tente de rendre compte de l'évolution chronologique de la technologie et son adoption dans la finance. Les systèmes experts sont apparus dès les années 1980, notamment dans les banques et les compagnies d'assurance. Les premières applications dans le Web sémantique et l'interaction « naturelle » avec les utilisateurs remontent, elles, à la fin des années 1990.

Il convient dans cette logique de relier l'émergence actuelle de l'IA au big data et de dissocier les technologies qui font appel aux techniques traditionnelles du data mining, du clustering et de l'analyse de composants de celle qui font appel à des données non structurées (deep learning, réseaux neuronaux…). L'engouement actuel pour l'IA se concentre principalement sur ces dernières, et particulièrement l'apprentissage profond.

A ce jour, les utilisations de l'IA dans la finance relèvent de l'application, par des automates, d'algorithmes de décisions et/ou d'apprentissage pour résoudre des problèmes dans un cadre bien défini. Ainsi par exemple :

- des algorithmes de machine learning qui calibrent un modèle sur des données historiques pour déterminer la décision optimale dans les situations futures ;
- des processus d'allocation dans une stratégie d'investissement qui améliore le portefeuille en se basant sur une fonction de « regret déterministe ».

Ces deux domaines, les robo-advisors et le trading automatique, semblent destinés à être les premiers où l'IA se diffusera massivement. Bien que la confidentialité soit jalousement gardée, l'intelligence artificielle est probablement assez largement répandue dans le trading algorithmique et pourrait devenir incontournable sous la forme d'outils d'aide à la décision. Tout autant que l'automatisation des processus de trading, les algorithmes de machine learning promettent en effet des gains d'efficacité et de performance ainsi qu'une réduction du risque opérationnel. Certaines start-up offrent déjà des outils d'analyse complexe, permettant de diminuer les prises de risque ou d'analyser de grandes quantités de données externes.

Quant aux robo-advisors, l'idée principale est de créer des conseillers virtuels capables, à l'aide de questions posées au client, de cerner son profil de gain et de risque, pour lui proposer un produit d'investissement adapté. Pour le moment, les produits semblent en général plus proches d'une interface marketing liée à un algorithme de matching qu'à des algorithmes d'intelligence artificielle. Mais le besoin est réel, et le gain potentiel important. En France, le premier robo-advisor auto-apprenant vient d'être labellisé par le pôle Finance Innovation.

Même si les algorithmes évoluent rapidement, les véritables ruptures paraissent dériver de la puissance de calcul et des données :

- Pour la première, la disponibilité des GPU (*graphics processing units*) dans les data center lève le principal obstacle à la diffusion d'algorithmes d'IA très gourmands en calcul ;
- Pour les deuxièmes :
  - la quantité de données disponibles double tous les deux ans ;
  - les données aujourd'hui incluent des informations complexes comme des images, des vidéos, du texte, des liens humains… ;
  - les acteurs, en particulier en finance, prennent conscience de la valeur de ces données (cf. note de cadrage en annexe).

On passe ainsi d'une course à la vitesse à une course à la finesse. La vitesse n'est plus le centre d'intérêt qu'elle a été en finance quantitative. Ce que permet l'IA aujourd'hui est d'accéder à des signaux et motifs plus subtils, au sein de données plus complexes et en expansion permanente, et de les exploiter pour prendre des décisions mieux informées – ce qui est l'essence de la finance.

Les questions qui se posent dès lors sont de quatre ordres :

1. Estimation du risque et credit scoring :
   a. Octroi de crédit par l'utilisation du graphe social (par exemple contacts sur Facebook ou LinkedIn comme estimateur du risque de crédit efficace) ;
   b. Domination des grandes plates-formes américaines (GAFA) sur ces données individuelles / personnelles.

2. Analyse financière et recherche, investissement et stratégie de portefeuille, trading :
   a. Analyse de sentiments, de textes et de signaux faibles ;
   b. Les robo-advisors semblent encore généralement rudimentaires, mais prometteurs ;
   c. Annonces de taille :
      i. Blackrock communique sur son utilisation de machine learning et IA pour le Forex et l'equity,
      ii. Virtu Financial communique sur son utilisation massive du big data depuis des années, et affiche un seul jour de trading négatif en 6 ans entre 2009 et 2014.

3. Fraude, lutte anti-blanchiment et contre le financement du terrorisme, KYC, compliance :

a. La détection de fraude est un des exemples de succès du machine learning et de l'IA, déjà appliqué pour l'octroi de crédit et l'assurance ;
b. La manipulation du LIBOR a été détectée en premier lieu via des analyses statistiques.

4. Robotisation de tâches non spécifiques à la finance :
    a. Service client (chatbots) ;
    b. Marketing (offres personnalisées, gestion de l'attrition) ;
    c. Comptabilité et direction financière.

L'entrée de nouveaux acteurs (Fintech) sur des services habituellement réservés aux banques ou assurances dépendra en large partie des réponses, technologiques et de marché, mais aussi réglementaires, à ces questions.

Un dernier débat porte sur le pas à franchir pour la machine afin de devenir un être indépendamment intelligent, capable de développer un comportement intelligent et non pas seulement d'en produire un conçu par son programmeur. Celui-ci semble pour l'heure assez abstrait. Puisque l'humain ne peut déjà plus nécessairement suivre la complexité de la prise de décision de la machine, il s'agit alors déjà de comportements intelligents non prévus.

On ne peut en tout cas que spéculer sur l'effet de ces comportements au fur et à mesure que l'IA se diffuse dans le système financier et que les problèmes autour de l'apprentissage automatique (« unsupervised training ») sont résolus. Mais cette dimension doit être intégrée à la réflexion.

II Risques

En matière financière, bancaire et assurantielle, les questionnements soulevés par l'IA relèvent de cinq ordres au moins :

1. Automatisation et responsabilité :
    a. Au fur et à mesure que plus en plus de tâches et de décisions seront déléguées aux algorithmes, la notion de responsabilité se brouille ;
    b. La Securities Exchange Commission (SEC) a récemment émis un avis rattachant la responsabilité sur le comportement des roboadvisors.

2. Auditabilité :
    a. Certains algorithmes sont opaques, et leur output difficile à interpréter, rendant la compréhension d'une décision algorithmique parfois impossible ;
    b. Cela pourrait rendre plus complexes à détecter les délits d'initiés ou les abus de marché par exemple ;
    c. Il convient d'examiner également le processus de décision des IA : quels éléments sont pris en compte ?

3. Prédictibilité et fiabilité :

a. Comme tout programme informatique, une IA est sujette aux bogues, au piratage et aux erreurs d'implémentation. Une IA peut aussi être mal conçue et comporter des erreurs intrinsèques, ou des "angles morts" face à des situations inédites ;
b. Il est déjà parfois attribué aux programmes de trading automatique la capacité de perturber les marchés, sans que cela ne soit nécessairement la seule cause : flash crash de mai 2010 sur le Nasdaq, « perturbation » des logiciels de Knight Capital Group en août 2012…
c. Les risques de déploiement de logiciels insuffisamment testés semblent destinés à augmenter.

4. Assurabilité et risque systémique :
    a. La diffusion de l'IA fait naître de nouveaux risques et modifie les grands domaines d'expertise de l'assurance (e-santé, véhicules autonomes…) ;
    b. L'assurabilité de ces nouveaux risques reste à construire, en fonction des contraintes ci-dessus ;
    c. En revanche, en l'état de la diffusion de l'IA sur les marchés financiers et bancaires, l'hypothèse d'un nouveau risque systémique semble lointaine : le risque principal que pose l'IA est celui d'un manque de maîtrise et d'une mauvaise utilisation, ou d'un jeu de données test truqué, plutôt que d'algorithmes transients décidant de prendre leurs libertés ;
    d. Les marchés financiers pourraient en revanche faire l'objet d'une cyberattaque spécifique, voire généralisée, comme celle dont a fait l'objet l'Estonie dès 2007. Tout comme dans bien d'autres domaines, le déploiement de l'IA posera des questions nouvelles en termes de cybersécurité.

5. Discrimination et vie privée :
    a. La disponibilité et l'utilisation des données, et notamment des données personnelles, est au cœur du développement de l'IA. En matière financière, les réglementations LAB-FT se croisent donc avec les réglementations CNIL et les directives européennes. Ce point est développé ci-dessous.
    b. L'utilisation de signaux faibles couplé à l'opacité des décisions algorithmiques pourrait mener à de nouvelles formes de discrimination, compliquées à identifier et donc à combattre – en tout cas tôt.

III Enjeux réglementaires

Ainsi qu'il a été évoqué ci-dessus, la sécurité devrait devenir un sujet encore plus important pour les banques, les compagnies d'assurance, les places financières et leurs régulateurs. Cependant, l'IA offre également des perspectives importantes dans la détection de fraude / compliance. Par exemple, la start-up américaine Neurensic offre, grâce à ses algorithmes, de pouvoir détecter en temps réel des comportements frauduleux sur les marchés financiers. La capacité de pouvoir détecter automatiquement ces incidents permettrait d'accroître la sécurité des marchés, limitant ainsi les possibilités de manipulation des cours. Plus généralement, la détection de fraude, de délit d'initié et d'abus de marché devraient se trouver facilitées par un audit plus exhaustif et rapide grâce aux données et à l'IA.

Dans cette perspective, des initiatives comme celle de la SEC de créer un entrepôt de données unique sur toutes les transactions qui se déroulent sous sa surveillance[1] pourraient avoir un impact majeur.

Le développement des algorithmes d'IA, en effet, dépend largement de la quantité de données disponibles. Et dans ce contexte, la France fait assez pâle figure face aux Etats-Unis et à la Chine. Confrontées à ce manque de données accessibles, les start-up françaises évoluent dans un contexte défavorable.

Afin de changer d'échelle, plusieurs pistes doivent être explorées, aussi bien à l'échelon national qu'européen :

- Une concertation plus fluide entre régulateurs, institutions de place et fintechs ;
- Une gouvernance des données qui offrirait un cadre plus clair et balisé pour créer un terrain favorable à l'innovation, surmontant les difficultés réglementaires auxquelles se heurtent aujourd'hui les institutions financières ;
- La poursuite des efforts de mise à disposition des données publiques (open data) ;
- Une incitation à mettre en place des plates-formes anonymisées en open data ;
- Une telle incitation peut être l'occasion de donner une prime aux entreprises qui donnent du pouvoir au consommateur sur ses données.

La diffusion de l'IA pourrait donc rendre nécessaire d'aller plus loin que la loi République Numérique et le règlement européen GDPR (General Data Protection Regulation).

La question de l'audit des logiciels se posera également avec une acuité croissante au fil du développement de l'IA. Prendront un relief de plus en plus important :

- la propriété intellectuelle sur le logiciel ;
- la capacité à accéder au code source des algorithmes.

Les décisions prises en la matière permettront, le moment venu, de constituer un comité d'éthique ou de renforcer la CNIL afin de faire face au risque de nouvelles discriminations.

[1] https://www.sec.gov/news/pressrelease/2016-240.html

Conclusion

Les banques, les compagnies d'assurances, la place financière et les fintechs se mesurent simultanément à quatre enjeux technologiques :

- l'intelligence artificielle, en passe de transformer leurs métiers du service client jusqu'à l'investissement ;
- la Blockchain et les technologies de bases de données distribuées ;
- les processus d'automatisation robotisés (RPA) ;
- la cybersécurité.

Pour ce qui est de la première, trois réflexions sont à mener, en matière de données, de propriété intellectuelle et de sécurité. La création d'un groupe pluridisciplinaire, représentatif de la place financière, bancaire et assurantielle, est à envisager, avec pour fonction d'assurer une veille technologique et instaurer un dialogue fluide avec les régulateurs. Ce groupe pourra évoluer, le cas échéant, vers la création de consortiums suivant l'exemple de Blockchain (cf. encadré en annexe).

Recommandations

1/ Améliorer le partage des données, selon des configurations efficaces et respectueuses de leur valeur.

Les sociétés en IA appliquée à la finance ont peu accès aux données, ou accès à peu de données. La qualité des travaux s'en ressent alors. Il convient de réfléchir plus avant aux modalités d'ouverture et d'élargissement de ces données, en expertisant les éventuels freins réglementaires injustifiés et les moyens de les lever, ainsi qu'en étudiant les mécanismes de gouvernance permettant de rendre l'analyse coût / bénéfice positive pour chacun des acteurs concernés.

Sur un autre aspect, les mêmes données peuvent se retrouver à l'identique chez plusieurs acteurs, notamment des données en matière de KYC (Know Your Customer). Des initiatives sectorielles en ce sens pourraient être promues.

Le PIA 3, s'agissant d'applications à la finance, pourrait chercher à privilégier dans les appels à projets ce type d'initiative (logique de consortium ou assimilé).

2/ Identifier plus avant les enjeux de régulation.

Il paraît nécessaire de pouvoir anticiper autant que possible du côté des autorités de régulation et de supervision les enjeux qui émergeront de l'application de l'IA en finance, afin de promouvoir les réponses les plus efficientes possibles, à la demande le cas échéant des acteurs de marché ou des consommateurs. Un cadre juridique clair et précis est en effet un vecteur rassurant, voire incitatif, pour engager des développements bénéfiques et pérennes. En même temps, il est certain que le cadre juridique doit aussi pouvoir évoluer continument, au gré des avancées technologiques et de leur industrialisation. Enfin, il est à noter que les autorités de supervision elles-mêmes peuvent bénéficier des progrès de l'IA ; elles ont aussi, notamment, des masses de données à traiter intelligemment.

L'inclusion des autorités de supervision dans la réflexion collective apparaît indispensable, sous une forme devant leur permettre de jouer un rôle équilibré.

3/ Améliorer la connaissance collective sur les différentes initiatives, afin d'identifier les activités / les segments les plus susceptibles d'innovation rapide / d'innovation disruptive.

Les deux recommandations précédentes doivent permettre de nourrir un dialogue continu, fondé sur la pratique et non sur la théorie, et mutuellement bénéfique entre autorités de supervision d'un côté et acteurs de marché de l'autre. Un modèle possible est celui qui a été mis en place – même s'il doit lui aussi être continument amélioré – pour les FinTech ou sur ce qui est fait vis-à-vis de la Blockchain (ouverture des régulateurs afin d'inciter les acteurs de marché à développer les expérimentations dans un cadre maîtrisé, possibilité d'enveloppes du PIA 3).

S'agissant le cas échéant de travaux particulièrement confidentiels, les autorités publiques et notamment de supervision peuvent en tout cas jouer un rôle de « tiers de confiance »

précieux pour avancer dans la compréhension de place des enjeux et potentialités. Le mapping selon les deux axes rapidité d'industrialisation / potentiel de disruption, dont le groupe de travail a commencé une ébauche, ne peut être poursuivi que dans un cadre robuste d'échanges. Un tel travail pourra également servir le cas échéant au PIA 3.

ANNEXES

Membres du groupe

Rapporteurs :

- **Sébastien Raspiller**, Direction Générale du Trésor, sous-directeur Financement des entreprises et du marché financier
- **Jean Rognetta**, président, PME*finance*-Europe*Entrepreneurs*

Membres :

- **Marouane Anane**, BNP Paribas
- **Guillaume-Olivier Doré**, Robin'Finance
- **Joëlle Durieux**, pôle Finance Innovation
- **Jean Ferré**, BCG
- **Julien Murésianu**, Jalgos
- **Maximilien Nayaradou**, pôle Finance Innovation
- **Nicolas Rivard**, Euronext
- **Milie Taing**, Lili.ai
- **Pierre-Alexis de Vauplane**, PME*finance*-Europe*Entrepreneurs*

Cadrage initial

Le secteur financier se caractérise notamment par deux aspects :

- L'importance de son rôle de « tiers de confiance » basée en particulier, s'agissant de la relation clients, sur le rôle humain – l'intermédiaire financier disposant d'une connaissance parfois intime, car ayant trait à des aspects financiers, de ses clients ;
- une régulation plus fournie et une supervision plus intrusive qu'ailleurs, imposant et requérant de la part du secteur financier la fourniture et la production de multiples données.

Le secteur financier a de longue date investi dans le traitement fin des données dont il dispose ou dans l'élaboration des données qu'il doit produire. Cela étant, des investissements informatiques pourtant lourds n'ont pas toujours permis aux entités du secteur financier de disposer en leur sein des systèmes les plus en pointe et interopérables entre eux. Face aux progrès récents et massifs dans les méthodes de traitement fin des données, le secteur financier investit encore davantage sur le sujet, en partenariat le cas échéant avec des start-up.

Le développement des techniques de big data est ainsi à même de transformer profondément le secteur financier sur de nombreux aspects. L'amélioration continue des techniques d'intelligence artificielle trouve ici, au vu de la masse de données recueillies et à fournir, un terrain de jeu potentiel : il convient déjà de lister les segments d'activité les plus concernés, et d'actualiser cette liste très régulièrement.

Le groupe de travail a vocation à se concentrer sur ce que le couplage big data / intelligence artificielle pourrait impacter parmi les activités financières, selon les deux axes suivants au moins à ce stade : i) fonctionnement des marchés financiers (microstructure de marché, et notamment trading) ; ii) relation clients et impact sur le matching entre l'offre et la demande.

Il apparaît utile en premier lieu de bien circonscrire le sujet. Il est en effet beaucoup question depuis quelques années des FinTech, ce qui peut recouvrir un très large spectre allant du crowdfunding à la blockchain – sans que l'intelligence artificielle n'ait à voir avec cela. De manière générale, il y a une tendance globale qui touche aussi le secteur financier à rechercher des processus plus efficients par une remise à plat profonde et rapide de l'utilisation des données et des informations. En d'autres termes, les activités financières sont également concernées - comme beaucoup d'autres - par la numérisation de l'économie. Il est possible par exemple de constater la progression importante de la bancarisation en Afrique du fait de l'essor de la diffusion des réseaux mobiles.

En revanche, l'utilisation de l'intelligence artificielle dans le secteur financier en est à ses prémices, même si l'engouement pour l'IA remonte au début des années 1980 et aux premiers systèmes experts. Ceux-ci se sont révélés plus dispendieux et difficiles à maintenir que prévu et n'ont pas causé l'impact dans les services financiers que certains pouvaient attendre d'eux. L'engouement lancé par le succès de « l'intelligence machine » symbolisée par des projets comme IBM Watson et par les victoires aux échecs et au go doit encore donner des fruits.

Et ce, d'autant plus si l'on considère une définition de l'intelligence artificielle nécessitant qu'il y ait apprentissage ou interaction. Sur cette définition, les robo-advisors actuels ne font pas nécessairement appel à l'intelligence artificielle, s'il s'agit de produire un schéma d'allocation des actifs en fonction seulement de données initiales fournies par le client (i.e. s'il y a seulement exécution automatisée).

Pour autant, il est à attendre que le développement d'applications pratiques de l'intelligence artificielle se développe de plus en plus dans le secteur financier, à un horizon qui reste difficile à estimer. Le déclin du coût du calcul et la grande quantité de données disponibles sur les réseaux mobiles et sociaux créent des conditions favorables à l'intelligence artificielle. Et comme indiqué *supra*, le secteur financier offrant des caractéristiques propices à une application rapide des progrès en matière de traitement fin des données, il est possible d'estimer que si les activités financières ne seront pas nécessairement les premières impactées, elles seront aussi loin d'être les dernières.

De fait, les principaux déploiements actuels ou dans un avenir proche, tels que l'on peut les prévoir à ce stade, pourront concerner : la gestion de portefeuille (allocation par les robo-advisors), le trading algorithmique, la détection des fraudes... En termes de domaines d'application prévisibles, cela correspond au service à la clientèle (chatbots et interfaces de conversation), à la sécurité (notamment des transactions), à l'analyse des informations et des données – en élargissant vers des données moins comptables et formelles – et à la vente et au conseil en matière de vente de produits financiers.

A moyen terme, la conjonction des données personnelles, comportementales et financières/bancaires/assurantielles laisse entrevoir l'émergence de possibles assistants financiers personnels, inspirés des services Amazon Echo et Google Home, et de véritables services financiers personnalisés, par exemple par le biais d'agents intelligents optimisant les besoins des assurés et pilotant la gestion de leur épargne.

La capacité des épargnants et des consommateurs à distinguer les divers services d'intelligence artificielle et à en apprécier les qualités respectives est inconnue. Son évolution influencera son acceptabilité ainsi que l'émergence d'un oligopole restreint, comme dans les plates-formes numériques aujourd'hui, ou la préservation d'un marché concurrentiel.

En second lieu, il importe – notamment du fait que les activités financières sont hautement régulées – de pouvoir apprécier rapidement les éléments potentiellement positifs et les éléments potentiellement négatifs que le développement de l'intelligence artificielle peut amener.

Parmi les éléments potentiellement positifs s'inscrivent déjà ceux mus par la recherche d'un meilleur service aux clients ou d'une meilleure efficience des procédés. Ces éléments doivent naturellement émerger de la pression concurrentielle existant au sein du secteur financier, par une démarcation positive de l'offre ou par une meilleure maîtrise des coûts et in fine des prix. Il est également possible d'imaginer qu'une appréciation plus granulaire des clients permette de satisfaire davantage de besoins qu'aujourd'hui, sans nécessairement accroître les risques pris.

Toutefois, des éléments potentiellement négatifs peuvent aussi apparaître, en fonction notamment de l'utilisation recherchée : des techniques de trading automatisées « intelligentes » peuvent-elle être conçues dans une logique d'abuser le marché sans être « vues » ? Pour prendre un autre exemple : un robo advisor « intelligent » peut-il apporter un élément positif en étant programmé pour ne servir que les intérêts propres du client – sans autre motivation ? ou peut-il être conçu pour maximiser les frais perçus ?

A plus long terme, suivant les craintes de certains scientifiques de renom, comment éviter – en termes de régulation et de supervision – que les programmes d'intelligence artificielle puissent être susceptibles d'évoluer hors contrôle et de manière non bénéfique au public ? Comment s'assurer que ces programmes d'IA soient détectés et isolés immédiatement, dans un contexte où ils interagissent en temps réel avec d'autres programmes financiers ?

Au-delà de ces enjeux qui ne sont pas nécessairement propres aux activités financières, il convient aussi de se pencher sur quelque chose qui est en revanche propre au secteur financier : sa nature potentiellement systémique, en matière de stabilité financière. Cet aspect nécessite que les développements à venir puissent être rapidement et correctement appréciés, afin le cas échéant de s'en servir pour renforcer la stabilité financière ou au contraire de les corriger rapidement voire de les circonscrire pour éviter de menacer la stabilité financière. Cela étant, et pour rappel, aucun enjeu de ce type n'apparaît à ce stade aujourd'hui.

Le principal enjeu du groupe de travail, au vu des premiers développements exposés dans la partie précédente, consiste à définir un premier cadre pour un dialogue continu entre d'un côté les entités du secteur financier faisant appel à l'intelligence artificielle et de l'autre côté les régulateurs et superviseurs, et ce dans un double objet d'une part d'attractivité et de compétitive et d'autre part de bonne maîtrise des risques financiers.

A ce stade au moins, il n'y a pas de raison de penser que l'on ne pourrait pas reproduire le premier cadre d'un tel dialogue tel qu'il a été défini pour la blockchain.

Il convient ainsi de pouvoir identifier, collectivement, les activités financières où l'arrivée de l'intelligence artificielle peut être rapide et/ou avoir assez vite des impacts profonds. Pour la puissance publique, il ne s'agit pas seulement d'avoir une vision passive en suivant les développements mus par le jeu des affaires, mais aussi d'adopter une vision active pour disposer le cas échéant d'une capacité d'adaptation ou de réaction rapide et de peser dans les débats européens ou internationaux.

Il s'agit donc d'une veille active, permettant la prise de conscience de tous les acteurs concernés, et de pouvoir en tirer profit plutôt que de subir des standards imposés ailleurs. Ce ne peut être que le début d'un dialogue continu, sur la base tout d'abord d'une estimation partagée de l'impact potentiel de l'intelligence artificielle sur diverses activités. Par exemple, l'intelligence artificielle va-t-elle faire évoluer les techniques de trading, et dans ce cas comment s'assurer que les transactions sont toujours efficacement supervisées ?

De manière générale, ces échanges devraient s'inscrire selon quelques lignes directrices, qui pourraient être les suivantes sous réserve des conclusions auxquelles aboutira le groupe de travail :

- dialogue et accueil des nouveaux entrants et/ou des impétrants développant de nouveaux modèles ;
- attitude proactive partagée d'identification des points de frottements éventuels entre nouveaux modèles et cadres de régulation et de supervision existants, et recherche des solutions appropriées ;
- compréhension des risques et avantages associés aux nouveaux modèles pour anticiper leurs effets micro- et macroéconomiques potentiels, tout en restant neutre quant aux choix commerciaux et technologiques effectués.

Blockchain

Qu'est-ce que la blockchain – quelles sont ses applications potentielles en matière de transactions sur titres ?

Les technologies de registres distribués, souvent appelées « blockchain », permettent aux participants d'un réseau de valider des opérations entre eux, sans l'intervention d'un tiers à l'opération qui serait chargé d'effectuer ce contrôle. Cette procédure de validation repose sur des mécanismes de consensus. Le registre distribué est accessible à tous les participants du réseau et contient l'historique de l'ensemble des opérations. La blockchain peut présenter un intérêt dans de nombreux domaines. Pour les opérations sur titres, elle a principalement deux avantages : i) permettre d'opérer rapidement et efficacement le transfert de propriété du titre d'une partie prenante à une autre ; ii) enregistrer tout au long de la vie d'un titre chacune des opérations de cession dont il a fait l'objet, dans un « bloc » qui vient s'ajouter aux blocs précédents – cela permet de retracer de manière très transparente le « cheminement » du titre depuis son émission.

Les initiatives blockchain en France en matière financière - deux textes à venir :

Un décret autorisé par l'ordonnance relative aux bons de caisse pourra préciser dans quelles conditions les « minibons », des instruments de dettes de PME qui ne sont pas des instruments financiers au sens de la réglementation, peuvent être émis et transmis en utilisant la blockchain.
Une habilitation à prendre une ordonnance « blockchain » permettra, au cours de l'année 2017, d'adapter le droit positif afin d'autoriser l'émission et la transmission de certains titres financiers non admis aux opérations d'un dépositaire central de titres en utilisant la blockchain. Les titres concernés sont principalement les actions et obligations non cotées, ainsi le cas échéant que les parts de fonds et autres instruments de gré à gré.

Les objectifs poursuivis - l'industrialisation de l'activité de tenue de registre

Renforcer la fiabilité des registres de titres : les registres de titres non cotés sont généralement tenus par les émetteurs, souvent sous une forme « papier » peu fiable. Pour autant, cette activité est cruciale puisque c'est l'inscription en compte d'un titre financier qui matérialise sa propriété et le transfert de celui-ci. Il s'agit d'un sujet d'importance croissante avec notamment la montée en puissance du *crowdfunding*, qui conduit les entreprises à lever des fonds auprès de nombreux petits porteurs. L'utilisation d'une technologie de type blockchain pourra permettre d'achever l'adaptation du droit des titres à la dématérialisation, mutation complètement opérée pour les titres cotés, mais inachevée dans l'univers du non coté, afin de rendre les opérations plus rapides et plus fiable.

Encourager la souscription des titres, et faciliter, dans une certaine mesure, leur circulation : l'illiquidité des titres de capital non cotés est avant tout liée à des difficultés de valorisation. Cependant, les souscripteurs peuvent également être découragés par le formalisme et les frais d'inscription élevés qui sont exigés pour procéder à l'inscription de tels titres dans un compte-titre tenu par un établissement de crédit. De même, l'absence d'infrastructure facilitant leur transmission ultérieure constitue un frein important au développement de ces financements. La tenue de registres en blockchain pourrait permettre de lever ces obstacles : à terme, une telle évolution devrait simplifier le formalisme nécessaire pour inscrire des titres non cotés dans un compte-titre / PEA, et favoriser l'émergence d'un marché secondaire.

Renforcer la transparence et la traçabilité des opérations sur le non coté : La transparence inhérente aux opérations réalisées dans une blockchain (toutes les opérations sont connues et définitivement enregistrées dans un « bloc ») ouvre des opportunités intéressantes en termes de LAB-FT. La

réglementation à venir devra trouver un équilibre entre la protection de la vie privée et la confidentialité des opérations d'une part, et les impératifs LAB-FT d'autre part.

Et après ?

Lorsque la technologie aura prouvé sa validité sur des marchés circonscrits, tels que ceux des minibons ou du *crowdequity*, il est concevable qu'elle joue un rôle significatif pour simplifier et accélérer le traitement des opérations post-marché.

Sources

(liste non exhaustive)

1. Global state of Artificial Intelligence online survey, Forrester Research
2. Predictions 2017: Artificial Intelligence Will Drive The Insights Revolution, Forrester Research
3. Baker McKenzie, *Ghosts in the Machine: Artificial Intelligence, Risks and Regulation in Financial Markets*, avril 2016 http://www.bakermckenzie.com/-/media/files/insight/publications/2016/04/ghosts-in-the-machine/bk_fig_ghostsinthemachine_dec16.pdf?la=en
4. PWC : *Top financial services issues of 2017, Thriving in uncertain times*, décembre 2016 http://www.pwc.com/us/en/financial-services/research-institute/assets/pwc-top-financial-services-issues-2017.pdf
5. Bahrammirzaee, A., *A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems*, Neural Computer & Application (2010) 19: 1165
6. Duygun Fethi, M., Pasiouras, F., *Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey* ; European Journal of Operational Research, Volume 204, Issue 2, 16 July 2010

3. PRÉPARER UN CADRE FAVORISANT LE DÉVELOPPEMENT DE L'IA EN PRENANT EN COMPTE LES CONSIDÉRATIONS ÉCONOMIQUES, SOCIALES, ÉTHIQUES, DE SOUVERAINETÉ ET DE SÉCURITÉ NATIONALE

GROUPE DE TRAVAIL 3.1
SOUVERAINETÉ ET SÉCURITÉ NATIONALE

Groupe Souveraineté et sécurité nationale

Périmètre d'étude du groupe de travail

De manière générale, l'**Intelligence Artificielle** (IA) est un système (code, données et paramètres, interfaces, système physique, etc.) ayant une capacité à se comporter de manière adaptée dans un environnement plus ou moins complexe, notamment en élaborant, mémorisant et exploitant des représentations de cet environnement. On peut distinguer différentes formes d'intelligence selon la nature de l'environnement et des comportements considérés. Ces aptitudes particulières peuvent être précisément spécifiées et évaluées, s'il y a consensus sur la manière de déterminer le caractère adapté des comportements observés.

Au sens étymologique, la **souveraineté** désigne la capacité du souverain à maitriser les attributs dont il revendique avoir le contrôle : classiquement frontières, armée, police, monnaie, langue officielle, justice, etc. De manière plus générale, nous entendrons dans la suite de ce document **la souveraineté d'une entité donnée sur certains attributs comme étant la capacité de cette entité à maîtriser ces attributs dont elle revendique avoir le contrôle.** La souveraineté nationale historique en est le premier exemple. On parle aussi maintenant de souveraineté européenne, de souveraineté scientifique, de souveraineté technologique, de souveraineté économique, de souveraineté individuelle et, bien sûr, de souveraineté numérique.

La notion de **Sécurité nationale** désigne l'objectif de parer aux risques susceptibles de porter à une nation. En France, elle est définie (depuis la loi du 29 juillet 2009) par le nouvel article L.1111-1 du code de la Défense, qui dispose que « la stratégie de Sécurité nationale a pour objet d'identifier l'ensemble des menaces et des risques susceptibles d'affecter la vie de la Nation, notamment en ce qui concerne la protection de la population, l'intégrité du territoire et la permanence des institutions de la République, et de déterminer les réponses que les pouvoirs publics doivent y apporter ».

Le périmètre du groupe de travail 3.1 porte sur les souverainetés telle que décrites précédemment et la sécurité nationale. En effet, il est important de porter le débat autant au niveau de l'état (souveraineté et sécurité nationale), qu'au niveau de l'individu (protection de la personne et de ses données personnelles) et au niveau de l'entreprise (propriété industrielle).

Etat des lieux

Opportunités (points positifs)

- **Application de l'IA à la cybersécurité** : La majorité des fonctions de (cyber) sécurité classiques (chiffrement de données, authentification, contrôle d'accès, etc.) ne demandent pas d'intelligence artificielle, et ne devraient pas y avoir recours. Une caractéristique essentielle de ces fonctions de sécurité est en effet de faire ce qu'elles sont censées faire, et uniquement cela. Dans ce cadre, les techniques d'IA peuvent être utilisées pour **prouver** la correction et la complétude des programmes implémentant ces fonctions ou des protocoles utilisés. Dans ce cadre les techniques d'apprentissage peuvent par exemple permettre de construire les preuves de manière beaucoup plus efficace tout en étant aussi rigoureuses. Par ailleurs, certains domaines de la cybersécurité, émergeants ou non, peuvent néanmoins constituer des champs d'application pertinents pour l'intelligence artificielle, en particulier les systèmes ou algorithmes de fouille et d'analyse de « mégadonnées » et les mécanismes d'apprentissage. La détection d'activités suspectes ou malveillantes dans les réseaux, la reconnaissance et la caractérisation de logiciels malveillants, l'analyse de traces d'exécution,

l'analyse prédictive (par exemple, dans des SOC, Security Operating Centre), la recherche de compromission et le traitement d'incident, sont des exemples de ces domaines d'application.

- **Application de l'IA à l'extraction d'information ou de connaissance** : Traitement facilité des informations, capacité à traiter des grands volumes de données hétérogènes (images, vidéos, textes, données capteurs, etc.), l'une des applications directes, autant pour les entreprises que pour les services de l'Etat, étant l'open source intelligence (OSINT). Soulignons toutefois que le « travail » pour préparer les données avant la phase d'apprentissage par l'IA est fondamental. Citons par exemple l'opération d'annotation des images pour les applications de véhicules autonomes.
- **Application de l'IA pour la croissance et la compétitivité de l'économie française (souveraineté économique) :** Les récents progrès de l'Intelligence Artificielle (IA) ouvrent des perspectives de création de valeur et de croissance économique pour l'avenir de notre pays ; l'entreprise de conseil ACCENTURE publiant d'ailleurs une étude fin 2016 sous le titre : « Pourquoi l'Intelligence Artificielle est le futur de la croissance ? ». Pour le cabinet d'analyse TRACTICA, le marché de l'IA pour l'entreprise dépasserait les 11 milliards ($) en 2024. Ces technologies diffuseraient dans tous les secteurs de l'économie et contribueraient fortement à la croissance de la productivité des services et des industries (exemple de marchés importants en termes de e-commerce.)
- **Application de l'IA pour la transition environnementale :** le développement des réseaux intelligents (Smart Grid), indispensable pour l'introduction des énergies renouvelables non émettrices de CO2 et une meilleure allocation de l'offre et de la demande, requière les solutions d'échanges et d'exploitation de données entres les différents acteurs (particuliers, collectivités, Etat, industriels de la chaine) qui sont fortement dépendantes des technologies et systèmes à base d'IA.
- Rayonnement de la France.
  Que la France se positionne comme précurseur au niveau mondial aussi bien sur l'aspect diplomatie internationale (mise en place de définitions, de réponses en cas d'agression avérées, etc.) et sur l'aspect légal national (c'est à dire s'arroger les droits de saisir un serveur qui aurait une activité malveillante contre la France depuis l'étranger avec obligation de coopération du pays depuis lequel l'attaque a lieu).
- S'assurer que les entreprises qui font de l'IA en France pourront travailler de façon légale dans le pays afin de construire une propriété intellectuelle saine avec un environnement légal précis. Le risque c'est que si ce n'est pas fait, les sociétés décideront d'installer leur activité dans le pays qui l'accueillera avec un environnement légal adapté.

Menaces (points négatifs)

- Pour les algorithmes basés sur l'apprentissage machine (ML), une difficulté fondamentale est le phénomène de ``trou noir'', les données attirants les données. En effet les résultats issus d'un programme implantant une fonction seront d'autant meilleurs que le programme aura été entrainé sur une variété plus importante de données. On a donc intérêt dans une situation concurrentielle à soumettre un problème et les données spécifiques qui le concerne à l'entité dont le programme est déjà le mieux entraîné et par conséquent celui qui a bénéficié d'ors et déjà de plus de données. Ce faisant on bénéficie du programme a priori le plus performant, mais on lui offre en plus de s'entraîner sur ces nouvelles données… On a là une situation de conflit de valeurs cruciale aux enjeux éthique et de souverainetés :

favoriser la qualité d'une réponse en sacrifiant le cas échéant la souveraineté économique, nationale ou individuelle.

- Confiance aveugle dans les propositions du système sous couvert d'une décision algorithmique.
- Perte du contrôle des données utilisées par le système de l'IA.
- Difficulté à savoir ce qu'arrive réellement à faire une IA. Quelles sont ses capacités exactes ? Certaines IA sont-elles pus efficace/pertinentes que d'autres ?
- Automatiser et industrialiser les attaques en utilisant de l'IA pour les rendre adaptatives.
- Que le système légal ne suive pas les évolutions et que la France subisse sans pouvoir réagir.
- **Les biais des algorithmes et/ou des données**, notamment du fait d'une dépendance aux données, et donc les systèmes sont (très) sensibles à la qualité des données d'entrainement disponibles, dont rien n'assure la complétude peut provoquer **des biais sur les résultats**.
- **L'atteinte à la dignité humaine et au libre développement de la personnalité**.
- **Des décisions individuelles automatisées**.
- **Le profilage** facilité par l'apprentissage profond :
    - **Une évaluation de certains aspects personnels**;
    - **La possibilité de déduire des données beaucoup plus sensibles que les données brutes** d'origine.
- **Un renoncement à l'exercice de ses droits**.
- **Une banalisation de la collecte massive de données**.
- **Une difficulté d'application de certains principes de la loi « Informatique et libertés »**.
- **Risque à la souveraineté économique :** Une possible dépendance des acteurs économiques français à la surcouche IA développée par les quelques plus grands opérateurs, qui disposent des bases d'apprentissage les plus importantes et pertinentes, et par qui les données transitent forcément, ce qui leur donne un pouvoir considérable. Il s'en suivrait une captation de la valeur ajoutée produite par les acteurs disposant des systèmes et des technologies clés à base d'IA. Ce phénomène s'est produit dans le domaine de la culture et serait aujourd'hui susceptible de se généraliser à tous les secteurs d'activités. Une stratégie industrielle de filière (maîtrise de ces technologies par un, voire des acteurs français, et constitution d'écosystèmes) s'avère donc essentielle pour la souveraineté et la compétitivité de notre pays.
- **Risque à la sécurité nationale et industrielle : Un usage de l'IA par des sources malveillantes** : l'IA peut être utilisée comme moyens de recueil d'informations par des organisations criminelles (ingénierie sociale), et faciliter ainsi le vol direct ou indirect de données à caractère personnel ou industrielles. En outre, les vulnérabilités intrinsèques de l'IA (systèmes sur lesquels elle repose, et surtout leur fonctionnement) rendent des attaques nouvelles possibles (exemple : insérer des données erronées dans les antécédents d'apprentissage, pour faire dévier le résultat du système, en introduisant donc des biais volontaires et calculés dans les données).
- **Risque de souveraineté politique avec l'émergence de nouveaux acteurs sur la scène internationale.**
- **Une remise en cause de mesures jugées comme suffisantes** jusqu'à présent, du fait d'un accroissement des capacités (exemple : technologies d'identification automatisée), qui pourrait par exemple poser un problème de désanonymisation généralisée à l'avenir !
- Dans les menaces peut-être peut-on penser au sujet de la guerre de l'information et des techniques d'influence (Big Data à des fins électorales, analyse des sentiments sur les réseaux sociaux, etc.) même s'il faut distinguer ce qui relève de l'IA et du Big Data, mais on peut imaginer des algorithmes autonomes dans le champ médiatique.
- Séparer les risques liés à la sécurité nationale des risques liés aux OIV, par exemple une perte de contrôle de l'IA de gestion de la production d'électricité qui laisserait une région sans électricité.

Enjeux

1. Maîtriser et sécuriser les données
    a. Protéger les citoyens
        - Une frange importante des technologies issues de l'intelligence artificielle repose sur des techniques d'apprentissage et de deeplearning, lesquelles nécessitent l'accès à des bases de données. La qualité des résultats d'un système intelligent dépend en grande partie de celle des données utilisées au cours de la phase d'apprentissage. En fonction des domaines d'application, les données utilisées peuvent revêtir un caractère confidentiel ou personnel. C'est par exemple le cas de données médicales, dont l'exploitation à grande échelle pourrait permettre des avancées significatives dans le domaine de la pharmacologie, mais dont l'accès doit évidemment être réglementé et sécurisé pour garantir le secret médical, ainsi que la santé du patient potentiellement exposé à des cyber attaques. L'agrégation massive et l'exploitation par un tiers de données personnelles ou confidentielles, conjuguée à la perte de maîtrise par leurs « propriétaires » et l'opacité des traitements, constitue aussi un enjeu de sécurité important. On peut toutefois noter que ce sujet n'est pas non plus spécifique au domaine de l'intelligence artificielle.
        - Le développement de l'IA peut avoir une répercussion indirecte sur l'enjeu de protection des individus et citoyens. D'une part, les solutions développées peuvent être déclinées en une offre pour les particuliers. D'autre part, cela induit une sécurisation des produits vendus au grand public (exemple de la sécurisation de pacemaker vis-à-vis de cyber attaques).
    b. Permettre une grande maîtrise de ses données aux individus et aux organismes privés ou publics **(principe d' « autodétermination informationnelle »)**, pour qu'ils aient notamment la possibilité de décider eux-mêmes à qui et quand ils dévoileront des informations :
        - **L'intelligence artificielle (IA) doit être au service des capacités de développement personnel autonome, nécessaire au fonctionnement d'une société démocratique.** Le périmètre considéré, notamment pour couvrir la dimension éthique du thème, sous le prisme de la protection de la vie privée, est celui de la souveraineté « individuelle » : Voir articles de créations de la CNIL en mettant en avant les points suivants :
        - **Transparence** : diminuer l'asymétrie de pouvoir par une information préalable (sur l'existence du traitement, les objectifs poursuivis, les données collectées et leur circuit, la logique sous-jacente et les conséquences pour la personne), l'obtention du consentement spécifique (initial et dans le temps, face au carrousel de finalités ultérieures) et libre (détaché de l'obtention d'un service), ainsi que le droit de le retirer, de présenter ses observations, de contester les décisions et de s'y opposer (y compris un profilage), le droit de demander une intervention humaine.
        - **Améliorer la confiance envers l'IA** :
            - **Améliorer la phase d'apprentissage des algorithmes** : développer les systèmes adversaires pour améliorer les capacités de prédictions, contrôler régulièrement l'apprentissage ;

- **Développer les technologies protectrices de la vie privée**, notamment les solutions reposant sur la cryptographie (ex : homomorphie), Privacy by Design.

▪ **Donner la possibilité à l'individu d'utiliser ses propres données pour lui-même**, par exemple pour mieux se connaître, prendre de meilleures décisions... (cf. portabilité des données).

c. Avoir accès aux données

▪ Il faut aussi que la **collecte d'information ne soit pas vue par un pays étranger comme une attaque**. De la même manière qu'un moteur de recherche qui collecte des informations personnelles sur des français n'est pas une attaque et réciproquement. Mais c'est important qu'une personne puisse consulter ce que ce moteur collecte sur les français et qu'on puisse les faire retirer comme c'est le cas aujourd'hui.

▪ En plus d'informer les utilisateurs (individus et organisations privées ou publics), ceux-ci seraient eux aussi doter d'outils ou de technologies à base d'IA leur permettant de connaitre les informations extraites des données qu'ils fournissent. Ils pourraient alors réguler au travers de cloud personnel leurs comportements, voir s'assurer du chiffrement de données jugées sensibles.

2. Garantir la souveraineté, préserver et développer une capacité nationale et européenne d'action

a. Disposer des technologies et d'acteurs pour les exploiter de façon autonome

▪ La souveraineté dans le domaine de l'intelligence artificielle se traduit par la capacité à constituer et à soutenir une communauté d'experts et de scientifiques à même d'avoir une compréhension fine des systèmes à base d'IA et de leurs limites.

▪ Cet enjeu et les problématiques qui lui sont associées ne sont certes pas spécifiques au domaine de l'intelligence artificielle, mais il semble utile de le rappeler ici. La France doit donc se doter de moyens maitrisés **afin de maintenir et développer une capacité nationale d'action** dans la mise en œuvre des technologies et des systèmes à base d'intelligence artificielle. Une dépendance technologique à des acteurs extra-européens pour l'exploitation des données constituerait une menace importante à la confidentialité des données des citoyens, des entreprises et des services de l'Etat. En outre, nous assisterions à une perte importante de la valeur ajoutée produite sur le territoire, identique à ce qui s'est produit dans le domaine de la culture. La capacité à se doter de moyens maîtrisés constitue donc des opportunités économiques pour les acteurs industriels nationaux et européens

b. Pouvoir soit disposer de services équivalents, soit « protéger » les données

▪ La souveraineté dans le domaine de l'IA requiert également de développer et faire mûrir les technologies clés dans les organismes de recherche afin d'en irriguer le tissus industriel français, notamment de la défense, de la sécurité, etc.

▪ Ces services doivent couvrir l'intégralité de la chaine de la valeur, car les enjeux autour de l'IA portent également sur les réseaux (capteurs, smartphone, etc.) et les architectures « hardware » sur lesquels sont implémentés les algorithmes.

- Utiliser l'IA pour créer des structures de données fictives cohérentes et dynamiques utilisées dans le cyber-renseignement ;

c. Se prémunir d'attaques pour détourner la finalité des traitements réalisés par les systèmes intelligents
    - L'IA sert à **détecter des fuites de données, des failles de sécurité** (automatisation de la détection des vulnérabilités, automatisation des patchs).
    - Pour être compétitifs en termes de consommation énergétique, de compacité et de robustesse, les systèmes d'IA devront s'appuyer sur des architectures dédiées, y compris faisant appel au calcul analogique (ces architectures sont qualifiées de neuromorphiques quand elles sont inspirées du fonctionnement du cerveau). Ce rapprochement entre hardware et software permet aussi d'intégrer les données de manière spécifique à un système et de limiter les possibilités de duplication et donc le risque de fuite de ces données, ce qui est un enjeu de sécurité.

d. Garantir la transparence et la loyauté des traitements réalisés par les systèmes intelligents
    - Permettre dans un cadre éthique et légal la capitalisation et l'accès aux données ainsi que la capacité à les exploiter de façon autonome.
    - Il faut que la collecte de données soit légale pour que l'IA puisse faire son travail et s'épanouir. Et le fait que l'IA (et non un humain) travaille sur des données personnelles ne doit pas être un problème tant que les résultats en sortie ne contiennent pas de données personnelles.

e. Protéger les entreprises françaises et nos services publics, grâce à la constitution d'un écosystème d'acteurs sur la cybersécurité et à des mécanismes permettant l'accès à cette offre pour les PME /ETI
    - La protection des entreprises requiert la constitution d'un écosystème d'acteurs pour la cybersécurité et la compétitivité de leur technologie par l'apport de la recherche. Le tissu de la recherche français dans le domaine est important comme l'a montré l'analyse produite par ALLISTENE. Il s'agira donc de le soutenir, mais aussi de favoriser la maturation et la diffusions des technologies à des acteurs existants, voir nouveaux via la création de start-ups, dans le domaine de la cybersécurité.
    - Toutefois, l'accessibilité de ces solutions ou offres aux PME / ETI s'avère difficile par manque de financements et de compétences. Des mécanismes publics d'aide devront donc être mise en place pour soutenir la montée en gamme de notre tissus de PME / ETI sur ces sujets.
    - En effet, la souveraineté dans le domaine de l'IA repose sur le soutien des pouvoirs publics à la filière industrielle existante, voire à la création d'entreprises à même de devenir de véritables « champions » sur la scène internationale, sur des thématiques applicatives jugés prioritaires pour la souveraineté et la sécurité nationales tel que l'Open Source Intelligence (OSINT), la sécurité et la cybersécurité . (*troisième révolution stratégique, dite troisième offset, dans la stratégie de sauvegarde de la défense et de la souveraineté des États-Unis)*. Pour ce faire, la constitution d'écosystème compétitif, au même titre que ce que peut être la « Silicon Valley », est indispensable.

3. Tester, certifier et qualifier
    a. Amorcer une réflexion sur les méthodes et autorités compétentes pour tester et qualifier les IA qui serviront à des tâches critiques (Reco évaluation, métrologie, outils préalable pour la certification)
        - Pour pouvoir comparer dans un cadre commun les solutions de l'IA pour une tâche donnée, il faut de définir des cadres permettant de quantifier de manière reproductible, répétable et interprétable les solutions et systèmes (Data + algorithme + hardware) à base d'IA.
          Une première solution est la mise en place plus massive de challenges ou campagnes d'évaluation définie par les acteurs français. Les campagnes d'évaluation ou challenge sont des projets pluriannuels qui consistent à proposer un cadre commun de comparaison des solutions de systèmes. Une deuxième solution, évidemment complémentaire, pourrait être la mise en place d'une initiative nationale impliquant l'ensemble des acteurs concernés, qu'ils soient du monde de la recherche, de l'industrie, ou des organismes publics afin de développer des méthodes de métrologie et d'évaluations sociotechniques des systèmes à base d'IA. Pourrait être également adressé la problématique de la certification, point essentiel pour tous les acteurs industriels. Peu d'initiative existe sur le sujet, ce qui positionnerait la France en pointe dans ce domaine. Il est à noter que la majorité des campagnes d'évaluations sont aujourd'hui financées par les Etats-Unis en particulier par le DARPA et l'IARPA. Ceci a pour conséquence que les chercheurs et industriels du monde entier ont travaillé depuis plusieurs décennies sur les problématiques qui intéressaient avant tous les financeurs de ces campagnes. Faire que l'état français finance plus massivement ce type de campagne permettrait non seulement de créer un écosystème vertueux mais également de proposer des problématiques qui soient dans l'intérêt des entreprises et des concitoyens français.
4. Faire adopter l'IA dans l'administration
    a. Transformer la numérisation de l'état et passer du paradigme « formulaire papier vers formulaire électronique », à une notion de l'état au service de l'usager
    b. Par l'usage de l'IA pour fluidifier les rapports à l'administration
    c. Par l'automatisation des tâches répétitives, pour que l'agent puisse se concentrer sur son métier
    d. Par l'utilisation au sein de l'administration, qui développera l'écosystème de l'IA

Recommandations :

ConfIAnce : une plateforme de test, de qualification, voire de certification à la fin, pour les systèmes à base d'Intelligence Artificielle

Cette recommandation se décompose en deux volets. Le premier vise à mettre en place des challenges ou campagnes d'évaluation définie par les acteurs français. Le deuxième ambitionne de développer de nouvelles méthodes de métrologie et de validation, ainsi que des chaines d'outils pour la conception et l'intégration, adaptées aux systèmes à base d'Intelligence Artificielle. Cette initiative est évidemment essentielle pour la diffusion dans l'industrie, mais aussi pour la confiance et l'acceptabilité par les citoyens. Peu, voire pas d'initiative existe sur le sujet au niveau international ce qui positionnerait la France en pointe dans ce domaine. En outre, elle aurait un impact économique direct au travers la création d'entreprises ou de start-ups dans ce domaine. Enfin, elle fédérait l'ensemble des acteurs sociotechniques autour d'une initiative ambitieuse et pour un coût raisonnable.

FranceIA : une plateforme logicielle intégrative d'intelligence artificielle

Cette recommandation vise à fédérer, sous une impulsion forte de l'Etat, un écosystème national entre recherche et industrie en vue de développer une solution technologique nationale concurrente aux grandes solutions extra-européennes. Cette initiative permettrait de (1) ne pas dépendre des outils fournis par des acteurs étrangers pour l'exploitation de nos données, (2) d'éviter la perte d'une part croissante de la valeur ajoutée à leurs profits et (3) d'offrir une solution nationale d'intégration des briques technologiques développées en France.

Stocker les données

Cette recommandation vise à soutenir les solutions et entreprises nationales et européennes, notamment d'un point de vue géographique, de stockage et traitement de données. Pour ce faire, il s'agira d'accompagner au niveau européen le renforcement de la règlementation existante pour la protection des données personnelles, des entreprises et des services des Etats. La cryptographie homomorphe ou les techniques de cryptocalcul constituent une priorité et doivent être, à ce titre, davantage investiguées, notamment les technologies ou expérimentation préalable à un déploiement. Enfin, d'autres actions sont à souligner telles que la promotion de labels (exemple du référentiel déjà promu par l'ANSSI et le BSI « SecNumCloud »), le soutien à la demande notamment en réponse aux besoins de l'administration ou la promotion de l'utilisation de « webmail » personnel avec un stockage maîtrisé des données personnelles.

NEURO-IA : plateformes matérielles et logicielles autour des architectures neuromorphiques

Cette recommandation vise à soutenir le développement de plateformes matérielles et logicielles autour des architectures neuromorphiques. Pour se faire, il s'agira notamment de mettre en œuvre des plateformes, des chaînes d'outils ou des méthodes de conception, d'intégration, de validation ou de simulations adaptées à ces systèmes et composants microélectroniques. Ces développements constituent un enjeu majeur pour l'industrie française, notamment pour le futur des systèmes embarqués (ex : véhicule autonome, etc.), et nous permettrait de nous positionner rapidement sur le sujet de l'intelligence des systèmes décentralisés.

AutonomIA - apprentissage autonome fiable dans le temps

Cette recommandation vise à développer des techniques d'apprentissage automatique permettant à un système intelligent d'apprendre et progresser par interaction avec son environnement et ses utilisateurs sans avoir à transmettre les données correspondantes aux développeurs, et d'obtenir une performance optimale avec une supervision minimale. Il s'agit de combiner les meilleures techniques d'apprentissage (type deep learning) avec les approches de la robotique développementale (en s'appuyant sur des techniques de type "transfer learning, adversarial learning". Un impact majeur est d'interrompre le mouvement de concentration des données au profit d'un petit nombre d'acteurs qui fournissent les meilleurs services.

"PrIvAcy Compagnon" : un outil pour la transparence au service des citoyens et entreprises

Cette recommandation vise à développer puis diffuser de nouveaux outils (innovation, technologie à base d'Intelligence Artificielle) visant à réduire l'asymétrie d'information entre les citoyens, mais aussi les entreprises et les administrations, et les géants du numériques. Aujourd'hui les capacités d'exploitation de données sont détenues par les géants du numérique mais demain ces outils, disponible plus largement, permettraient d'extraire les connaissances (préférences, etc.) à partir de nos données fournies par les utilisateurs. Il serait alors plus aisé de réguler nos comportements, voire de s'assurer du chiffrement de données jugées plus sensibles. Ces solutions contribueraient à une dynamique d' « empowerment » sur internet.

Plateforme de renseignement et d'intelligence économique tels que les risques bancaires, etc.

Cette recommandation vise à développer puis mettre en place de plateforme de démonstration pouvant adresser le système dans sa globalité, c'est-à-dire allant de la collecte de flux de données (textes, images, sons, vidéos, etc.) à leur exploitation en passant par leurs mises en forme et leurs enrichissements. Les applications visées seraient notamment l'intelligence économique et le renseignement et serviraient autant les services de l'Etat que les entreprises. Notons également l'importance des usages et des méthodologies de recherche des utilisateurs. Toutefois, un échange en séance sur la légalité de ce type de solution. Il est rappelé à ce moment que ce type de solution existe mais qu'aucune n'est souveraine.

SecureIA : réseau de plateformes pour la sécurité et cybersécurité

Cette recommandation vise à mettre en place un réseau de plateformes de cybersécurité entre les territoires les plus dynamiques, d'un point de vue recherche et entreprises (la cartographie faite par le GT d'ALLISTENE serait susceptible de donner des indications sur ce point). Elle aurait pour objectif de favoriser la complémentarité des sites français, d'accompagner la croissance des start-up et des entreprises au travers de programme de recherche ambitieux, ainsi que de favoriser l'accès des PME / ETI à des offres et solutions technologies adaptées à leurs besoins ; par exemple par l'attribution de "voucher" technologiques.

Annexe 1 – Membres du groupe de travail

Laurent Voillot	DINSIC
Erwan Keraudy	Cybelangel
David Poincheval	ENS
David Naccache	ENS
Ingrid Valleaunu-paun	Interviewfrancophone
Julien Chiaroni	CEA
Jean-Gabriel Ganascia	LIP6
Stéphane Jourdain	LNE
Matthieu Grall	CNIL
Claude Kirchner	INRIA
Harold Ollivier	IMT
Jean-Luc Moullet	CGI
Edouard Geoffroy	ANR
Nathan De Lara	DGA
Thierry Goudon	DGRI

GROUPE DE TRAVAIL 3.2
IMPACTS ECONOMIQUES ET SOCIAUX

Anticiper les impacts économiques et sociaux de l'intelligence artificielle

Groupe de travail 3.2 – stratégie nationale en intelligence artificielle

Co-pilotes :
Rand HINDI (Conseil National du Numérique) et Lionel JANIN (France Stratégie)

Rapporteurs :
Charly Berthet (CNNum), Julia Charrié (France Stratégie), Anne-Charlotte Cornut (CNNum), Francois Levin (CNNum)

Cette note est issue des travaux communs de France Stratégie et du CNNum. Les documents en annexe ont été rédigés par les responsables des sous-groupes et n'engagent que leurs auteurs.

FRANCE STRATÉGIE
ÉVALUER. ANTICIPER. DÉBATTRE. PROPOSER.

CNNum
Conseil National du Numérique

Anticiper les impacts économiques et sociaux de l'intelligence artificielle

Groupe de travail 3.2, janvier-mars 2017

Analyse du groupe de travail

L'intelligence artificielle est un sujet "chaud" : elle bénéficie d'une dynamique, d'un momentum*, qui cristallise les attentions et les énergies. Il faut donc* ***agir maintenant*** *pour en tirer parti.*

Cette attention récente est fondée sur les progrès spectaculaires obtenus par la recherche en la matière. Il en résulte un accroissement brusque et imprévisible des tâches potentiellement automatisables qui amène à s'interroger sur les activités humaines. On passe d'un scénario de polarisation, dans lequel les emplois très manuels, d'une part, et les emplois basés sur les talents, d'autre part, semblaient préservés de l'automatisation, à un scénario de transformation qui touche potentiellement beaucoup plus de monde. C'est pourquoi la thématique de la disparition d'un grand nombre d'emplois tend à monopoliser le débat et à occulter tout raisonnement.

Le but de cette note est de sortir d'une vision simpliste (compréhension "magique" de l'IA, focalisation sur la disparition massive d'emplois, idée que l'IA nous tombe dessus quoi que l'on fasse induisant une passivité devant le choc annoncé). Une telle analyse a pour le moment empêché de mettre en avant les perspectives positives réelles (gains de productivité dans les entreprises, amélioration de la qualité des services) comme les alertes essentielles : danger de perte d'indépendance économique, absence de gouvernance des transformations, sous-estimation des impacts sur la vie quotidienne, en particulier dans le monde du travail, et nécessaire inventivité dans les accompagnements, en particulier par l'appareil de formation.

Ce document a une triple ambition :

1. Clarifier l'acception de l'intelligence artificielle et son intégration dans la production. Il s'agit principalement d'apprentissage machine – la capacité d'un logiciel informatique à reproduire avec une très bonne qualité certaines aptitudes humaines dès lors que des données d'entraînement en quantité suffisante sont disponibles. Ces dispositifs sont utilisés pour la reconnaissance de la voix, le classement d'images, de sons, de vidéo, le traitement de texte, par exemple la traduction automatique.

2. Attirer l'attention sur l'essentiel : les données. La valeur n'est pas là où l'on croit : elle provient des données nécessaires à l'apprentissage bien plus que de l'algorithme, dont les développements se font de manière ouverte (en open source*). En conséquence, les entreprises ou les administrations (pour la production de services publics) doivent se concentrer sur l'identification des données susceptibles d'être utilisées et sur les modalités concrètes d'exploitation (mise à disposition des données, éventuellement mutualisation, en conservant le contrôle des exploitations qui en sont faites ou en bénéficiant d'un juste retour), Par ailleurs les réflexions sur la libre-circulation des données, entre espaces géographiques mais également entre acteurs économiques doivent être approfondies.*

3. Souligner qu'il s'agira de transformations du travail beaucoup plus que de suppressions d'emplois : l'aménagement de ces transformations va demander une construction de la décision aussi anticipatrice que possible, incluant les acteurs pertinents, au premier chef les professionnels concernés. Le but à atteindre est de créer de la valeur dans le travail pour tous, de donner plus de pouvoir et d'intelligence grâce à l'IA et non de mécaniser les humains. Les politiques publiques, nationale et régionale, doivent être mobilisées pour construire une vision positive de l'IA, qui nécessite des plateformes plurielles d'évaluation et de co-conception des transformations.

Pour y arriver, un outil essentiel demeure : la ***formation****, qui ne doit pas être perçue comme un recyclage et un pis-aller. L'IA est un chantier de formation professionnelle essentiel qui peut servir de cas d'école pour les nouvelles visions de l'éducation et de la formation qui sont depuis plusieurs années débattues dans la société : importance des compétences transversales (*soft skills*), littératie numérique, à laquelle il faut adjoindre une littératie de l'IA. Une telle formation passe par une éducation informatique générale et une mobilisation de la société sur les formations initiales (entreprises, secteur de l'éducation populaire) mais aussi une inventivité dans les formations sur poste, les formations en conception, ce qui implique une forte mobilisation des entreprises elles-mêmes, de la recherche et des partenaires sociaux pour aboutir à de nouvelles voies.*

Si l'on ne fait rien, que se passera-t-il ?

- *au plan économique et diplomatique, il existe un risque de décrochage économique et de perte d'indépendance, si le phénomène de concentration de l'activité numérique autour de quelques grandes plateformes se prolonge ;*
- *au plan social : l'absence d'anticipation, la pauvreté des solutions d'accompagnement, des approches en silo et un pilotage assis sur la seule réduction des coûts auraient les effets dramatiques de destruction d'emploi que nous annoncent les cassandres.*

Il faut sortir des scénarios qui tétanisent toute prospective réaliste et plurielle, en écartant les scénarios irréalistes décrivant l'automatisation de toutes les fonctions productives. À cet égard,

le rapport ne se positionne pas sur la question du revenu universel, mais se concentre sur celle de la transformation des emplois, en vue d'une reconstruction de la valeur sociale des activités humaines.
Quel peut être le chemin ? Il est crucial d'organiser dès à présent un dispositif de concertation permettant d'anticiper ces impacts. Une grande plateforme numérique, alimentée par des débats dans les territoires, pourrait être envisagée avec tous les acteurs de l'emploi : Pôle emploi, l'AFPA, les acteurs régionaux, nationaux, européens, les syndicats. Elle pourrait ensuite être déclinée à différentes échelles : par branche, dans des réseaux thématiques pertinents, au niveau régional.

Pour ce qui concerne une politique inventive de formation – mobilisation de la formation initiale, des entreprises, de l'économie sociale et solidaire – elle suppose d'examiner les secteurs d'application de l'IA, en particulier d'affiner les modèles économiques pour les secteurs non industriels, comme l'éducation, la santé ou les politiques de la ville (smart city, en relation avec des politiques écologiques).

Le CNNum et France Stratégie soutiennent le besoin d'une telle concertation : il est dans leurs attributions respectives d'aider les politiques publiques à anticiper les impacts des transformations numériques et à en analyser les aspects combinés, économiques, juridiques et sociaux. Cette note vise à prendre date. Le calendrier restreint dans lequel se sont inscrits ces travaux ne permet, à l'heure actuelle, que de donner quelques éclairages.

Au-delà de cette sensibilisation, le CNNum entreprend de développer ses analyses sur la transition des emplois et d'approfondir ses propositions. Pour le CNNum, le chemin qui s'ouvre doit avoir pour but de préserver et d'augmenter la valeur du travail humain, et de concilier de façon innovante une économie compétitive et une amélioration du bien-être dans une société inclusive.

France Stratégie, pour sa part, entend continuer ses travaux d'animation du débat public sur l'IA, thème central de la révolution numérique, en donnant la priorité à l'emploi et à la dimension territoriale. En méthode, ces travaux seront menés en nouant toutes les collaborations utiles avec les administrations, les collectivités territoriales, les partenaires sociaux et les acteurs de la société civile.

INTRODUCTION

- **Historique** de l'intelligence artificielle et des ruptures récentes (*deep learning*) ; malgré des applications nouvelles à fort impact, nous sommes loin d'une IA "générale" : les logiciels actuels ne sont capables de réaliser que des tâches très spécifiques.
- Le déploiement actuel de l'IA s'inscrit **dans la continuité des transformations numériques à l'œuvre depuis des années** ; nombre des questions que celles-ci posent font déjà l'objet d'études par ailleurs (vie privée et gestion des données, éthique, acceptabilité et responsabilité, impacts écologiques tant en termes de flux et de stockage des données qu'en termes d'épuisement des ressources).
- **Les transformations à venir sont profondes, les incertitudes aussi.** Il s'agit d'une technologie au potentiel très significatif, capable d'étendre fortement le champ des possibles.
- **Tout n'est pas rose.** Il faut avoir à l'esprit certains points de vigilance : risques de perte de compétitivité, de dépendance et de renforcement de la concentration de la valeur sur certains segments de la chaîne de production.
- Un certain nombre d'études a attiré l'attention sur la destruction massive d'emplois. Sans se prononcer sur la pertinence des réflexions générales sur la redistribution de valeur, cette note se concentre davantage sur la **nécessaire transformation des métiers, notamment la complémentarité homme/machine**.

Reconnaissance d'images, filtre anti-spam à l'efficacité remarquable, système de recommandation de film ou de musique, traduction automatique, jeux vidéo aux personnages de plus en plus riches, effets spéciaux psychédéliques, victoire de la machine sur l'homme au jeu de Go, au poker il y a quelques semaines, agent conversationnel sur Internet ou à domicile et demain, logiciel de conduite automatique ouvrant la voie à la voiture sans chauffeur. Telles sont les réalisations permises par les progrès récents et spectaculaires de l'intelligence artificielle (IA), notamment les dispositifs d'apprentissage machine ou d'apprentissage statistique.

Si les principes de l'intelligence artificielle remontent loin, aux années 1940 par exemple pour les réseaux de neurones, un palier significatif a été franchi il y a quelques années, sous la triple poussée de l'augmentation de la puissance de calcul (permise en particulier par les cartes graphiques développées au départ pour le jeu vidéo), de la constitution d'immenses bases de données d'images, de vidéos ou de sons qualifiés (grâce à la volonté de quelques grands acteurs et aussi aux travaux des chercheurs et aux réseaux sociaux), et des progrès en algorithmique.

L'IA a certes fait des progrès spectaculaires, comme en attestent les exemples précédemment mentionnés et popularisés régulièrement dans les médias. Pour autant, les logiciels actuels ne réalisent que des tâches très spécifiques et nécessitent le plus souvent d'immenses volumes de données d'apprentissage pour aboutir à un résultat satisfaisant. Nous sommes encore très loin de l'intelligence artificielle dite

« générale » ou « forte », susceptible de prendre en compte le contexte et d'apprendre à partir de quelques exemples.

Continuités et discontinuités

Le déploiement de l'intelligence artificielle dans la société et l'économie en particulier s'inscrit plutôt dans la suite des transformations numériques à l'œuvre depuis des années. Il inclut aussi les progrès de la robotique, de la réalité virtuelle ou augmentée, et s'inscrit dans le prolongement de l'économie des données, dominée à l'échelle mondiale par les plateformes numériques américaines et asiatiques, avec un retard de la France et de l'Europe plus généralement.

L'IA soulève des questions classiques, qui ont déjà fait l'objet de nombreux travaux par ailleurs. À titre d'exemple, l'IA se nourrit de données pour faire des diagnostics personnalisés ; elle soulève donc des questions de données personnelles. L'enjeu est d'arriver à protéger la vie privée tout en tirant parti du potentiel de l'IA qui suppose d'avoir accès à ces données.

L'IA pose également des nouvelles questions éthiques et d'acceptabilité. Il y a une responsabilité du concepteur de l'algorithme mais également un impact des données utilisées pour réaliser l'apprentissage. Enfin, l'IA pose la question de la transparence des interactions humain-machine et de l'explicabilité des décisions que prennent les dispositifs automatiques. Refus d'un prêt bancaire, accident d'un véhicule autonome : il devra être possible de demander une explication sur la manière dont ont été prises les décisions, voire de déterminer des domaines où il est nécessaire que l'humain conserve la décision finale. C'est un sujet de recherche qui doit accompagner le déploiement de la technologie (cf. encart ci-dessous, résumant l'annexe 1 "L'intelligence artificielle en quête d'acceptabilité et de confiance").

Enjeux d'acceptabilité et de confiance dans l'intelligence artificielle
Sous-groupe Acceptabilité et de confiance

Le développement de l'IA est de nature à transformer considérablement la sphère de l'emploi et du travail, avec notamment pour conséquence la multiplication des systèmes d'IA dans l'environnement de travail, l'intensification des collaborations entre humains et machines ainsi que l'automatisation de la fonction de prédiction à travers les secteurs industriels et la délégation croissante à des machines de tâches et d'activités autrefois réalisées par les humains. Ces collaborations entre humains et machines vont revêtir toutes sortes de formes. Certains métiers pourront être entièrement remplacés par des machines autonomes (exemple : conducteurs) tout en créant des nouvelles tâches de supervision. D'autres resteront entièrement le propre de l'homme car le remplacement de l'homme par l'IA sera considéré inacceptable, inefficace ou non bénéfique pour la société. Dans d'autres cas encore, sans doute plus nombreux, des formes de coopération homme-machine verront le jour dans des environnements de travail où les IA seront des partenaires, des assistants plutôt que des remplaçants. Des agents autonomes participent déjà à différents types de négociations, de transactions (exemple : transactions boursières), d'interactions avec les clients, les usagers, les entreprises et organisations. À court terme, vont être affectés de manière encore plus large les métiers du calcul et ceux qui suivent une procédure formellement établie. De nouveaux métiers vont

également apparaître comme ceux de concepteur, d'entraîneur ou de gestionnaire (exemple : cogniticien) de données et d'informations, surveillant ou encore réparateur d'IA, etc.

Les membres du sous-groupe « Acceptabilité et confiance » ont mené une réflexion sur les concepts clefs dont la signification est susceptible d'évoluer sous l'effet de l'IA (responsabilité, mérite et valeur, créativité) et les risques de l'utilisation de l'IA dans certaines dimensions (affective, corporelle, mimétique, téléologique, créative) des métiers. Les effets positifs étant réels, gain de temps, prise en charge de tâches très répétitives ou des routines procédurales sans intérêt, il est néanmoins nécessaire d'envisager les risques, notamment ceux de désengagement, de déresponsabilisation, d'accroissement des disparités sociales, etc. Le résultat est un ensemble de recommandations qui concernent en particulier la nécessité d'expérimentations sur le terrain et d'études pluridisciplinaires. Un effort d'éducation/de communication sera crucial.

Pourquoi l'IA suscite-t-elle autant d'attention auprès du grand public et des médias ? S'agit-il d'une simple bulle amenée à se dégonfler ? L'expression "intelligence artificielle" possède un pouvoir d'évocation qui explique sans doute en partie les fantasmes qui y sont associés.

Transformations profondes, incertitudes importantes

Deux points à mettre en avant pour justifier l'attention particulière que mérite ce sujet de la part des pouvoirs publics, des entreprises et plus généralement de la société civile :
- d'une part, il s'agit bien d'une technologie au potentiel très significatif, capable d'étendre fortement le champ des possibles (jusqu'au véhicule sans chauffeur ou à la création automatique de scénarios, voire de films) ;
- d'autre part, il subsiste de fortes incertitudes sur son déploiement effectif et la manière dont elle sera appropriée, notamment dans le monde du travail, ainsi que sur la vitesse des progrès à venir. Le risque de destruction massive d'emplois écrase trop souvent toute considération sur l'intégration pratique de cette technologie.

Les questions du revenu universel, de la personnalité juridique des robots ou d'une éventuelle taxe sur les robots sont posées dans le débat politique[1]. Ce rapport ne se positionne pas sur leur pertinence mais précise le diagnostic, identifie des points de vigilance et formule quelques recommandations en envisageant plutôt des transformations du travail que sa disparition irréversible.

Enjeux juridiques de l'intelligence artificielle

L'IA et les bouleversements qu'elle est susceptible d'entraîner conduit à une réflexion sur les enjeux juridiques s'agissant notamment de la protection de l'IA ou encore de la pertinence des mécanismes de responsabilité actuels. Fort de ce constat, le sous-groupe « Enjeux juridiques et réglementaires de

[1] Rapport Delvaux adopté par le Parlement européen le 16/02/2017, Recommandations à la Commission concernant des règles de droit civil sur la robotique.

l'IA » a mené des réflexions et organisé des auditions sur le droit de l'IA afin de recueillir les observations des acteurs de l'IA, de la PME au groupe international, quant aux difficultés potentielles rencontrées en la matière ou au contraire, les solutions juridiques d'ores et déjà pratiquées de manière opérationnelle.

Les membres de ce sous-groupe s'accordent sur le fait qu'à ce stade de l'évolution de l'IA, il n'existe pas de nécessité impérieuse de repenser le dispositif législatif et règlementaire en vigueur. Les mécanismes et régimes juridiques actuels couplés à la souplesse contractuelle permettent aujourd'hui de faire face aux évolutions technologiques en apportant, sauf cas particuliers, un niveau de sécurité juridique satisfaisant. En outre, il ressort des réflexions menées que, comme sur d'autres sujets d'innovation de pointe, le droit ne doit pas devenir un frein à l'innovation. Aussi, la France et surtout l'Union européenne, qui apparaît dans tous les cas le niveau d'intervention adéquat pour éviter tout désavantage compétitif, doivent laisser le temps aux technologies et aux modèles économiques liées à l'IA de parvenir à une certaine maturité avant d'entreprendre le cas échéant un travail de complétion législative.

Le sous-groupe a réfléchi spécifiquement aux enjeux liés aux algorithmes, aux données, à la responsabilité des robots et à la protection des créations issues d'une IA. Ses réflexions sur ces thématiques sont annexées à la présente note (voir Annexe 2).

L'IA permet des gains de productivité au sein des organisations (filtre anti-spam, traitement automatique de dossier, réponse à des questions simples par un agent conversationnel automatisé[2]), des améliorations en termes de qualité de traitement (amélioration de la détection de tumeur[3]) mais elle n'échappe pas aux interrogations existantes sur les effets du numérique sur le bien-être ou sur la croissance économique (*cf.* débat entre les tenants d'un ralentissement durable des gains de productivité, la stagnation séculaire (*The Rise and Fall of American Growth*, Robert Gordon (2016) et les tenants d'une croissance à venir, *The Second Machine Age*, Brynjolfsson et McAfee (2014)). Pourquoi observe-t-on un ralentissement de la croissance de la productivité ? S'agit-il d'un manque d'investissement dans les technologies de l'information et de la communication (TIC), d'une insuffisante appropriation par les organisations, d'un déficit de formation ? Le numérique se contenterait-il de remplacer des processus existants et manquerait-il de nouveaux produits, comparables à ceux qui ont assuré la croissance économique lors des révolutions industrielles antérieures (avion, automobile) ? Sommes-nous dans une situation de croissance lente, durable ou s'agit-il d'un effet temporaire ? Le débat reste ouvert, au-delà de la question de l'IA[4].

---

[2] Logiciel capable d'interagir avec l'homme en langage naturel (par écrit ou par la voix) pour répondre à des questions simples.

[3] Lors du *Camelyon Grand Challenge for Metastatic Cancer Detection* qui s'est tenu en 2016, les meilleurs résultats ont été obtenus par les équipes faisant coopérer l'humain et l'IA : 0,5 % de taux d'erreur pour l'homme assisté de l'IA, 7,5 % pour l'IA seule, 3,5 % pour l'homme seul.

[4] Cf. la conférence organisée par la Banque de France et France Stratégie sur ce sujet le 01/02/2017 : *Productivité : une énigme française ?*

Ce qui est sûr, c'est que l'IA est une technologie générique (*general purpose technology* en anglais) : ses objets de travail (image, son, vidéo, texte, langage naturel) peuvent se retrouver dans de multiples activités, ouvrir des opportunités de créations de nouveaux produits et accroître les performances des utilisateurs. Les techniques d'apprentissage peuvent être diffusées d'un secteur à un autre, moyennant des adaptations limitées. Investir dans l'IA est donc susceptible d'avoir des effets d'entraînement dans de nombreux secteurs, à condition de réaliser des investissements complémentaires à cette technologie pour en tirer tout le potentiel. Ces investissements complémentaires sont à la fois matériels, en équipement numérique par exemple, ou immatériels, notamment en données, en logiciel, en savoir-faire et compétence ou en organisation. Les modèles économiques reposant sur le concept de technologie générique montrent en outre que les effets de substitution de la technologie au travail (les destructions d'emplois) peuvent être plus que compensés par l'amélioration de la compétitivité de certains secteurs, qui augmente leur potentiel à l'exportation (*cf.* exemple de l'Allemagne, très robotisée mais avec plus de salariés dans l'industrie qu'en France), par les gains de pouvoir d'achat, qui stimulent la consommation, et par l'innovation de produits, qui créent de nouveaux marchés, et donc de nouveaux emplois.

Points de vigilance

Pour autant, ce scénario rose doit bien entendu être nuancé. L'IA est l'une des pièces du débat sur le lien entre robotisation, automatisation et emplois, qui fait naître l'inquiétude d'une disparition massive et rapide de ces derniers. Sur ce sujet, les études, dont une revue a été faite notamment par le rapport publié en janvier 2017 par le Conseil d'orientation pour l'emploi[5] (COE), rappellent que les emplois qui sont automatisables sont en nombre limité (une dizaine de pourcents, que ce soit dans l'étude de l'OCDE[6], de France Stratégie[7] ou les travaux du COE). En outre, la possibilité technique d'automatiser des emplois ne veut pas dire qu'ils le seront, puisque ce résultat dépend de l'incitation économique à le faire et de choix d'organisation et sociaux. Les mêmes études établissent cependant que nombre de tâches sont susceptibles d'être transformées (près de la moitié). C'est sans doute la bonne lecture de l'étude d'Oxford (Frey et Osborne)[8] qui a contribué à l'attention portée au sujet.

Par ailleurs, l'IA est largement développée et commercialisée par les grandes plateformes du numérique (américaines et asiatiques), ce qui repose la question du retard français et européen dans le numérique et du risque de prolonger la dépendance actuelle. C'est parce qu'il existe un danger de se trouver à nouveau en retard dans la maîtrise et le déploiement de ces technologies que la stratégie nationale de l'intelligence artificielle est particulièrement bienvenue. Elle devra cependant *in fine* nourrir une stratégie européenne sur la question pour être totalement pertinente, comme pour nombre de questions numériques (fiscalité, données personnelles, etc.).

---

[5] Rapport du COE « Automatisation, numérisation et emploi », 10/01/2017.

[6] Arntz M., Gregory T. et Zierahn U. (2016), « The risk of automation for jobs in OECD Countries: A Comparative Analysis ».

[7] Le Ru N. (2016), « L'effet de l'automatisation sur l'emploi : ce qu'on sait et ce qu'on ignore », *La Note d'analyse* n°49, France Stratégie, juillet.

[8] Carl Benedikt Frey, Michael A. Osborne (2017), "The future of employment: How susceptible are jobs to computerisation?", *Technological Forecasting and Social Change*, volume 114, janvier, pp. 254–280.

Face à une telle transformation, l'enjeu de l'IA est donc celui de la diffusion de la technologie dans les processus de production et de la transformation des emplois.

Penser la complémentarité humain-machine

Une bonne intégration de l'IA passe par l'identification des apports que peut avoir cette technologie pour améliorer une organisation de travail, offrir de nouveaux services ou créer de nouveaux emplois, en assurant la mise en place d'une complémentarité entre l'homme et la machine. L'IA peut supprimer des tâches répétitives, pour permettre de se concentrer sur les plus intéressantes ; elle peut "augmenter" les travailleurs, en leur fournissant un outil de diagnostic qui enrichit leur expérience (réalité augmentée, pré-diagnostic en matière médical améliorant l'orientation des patients). Mais de tels outils comportent aussi le risque de transformer le travailleur en un simple exécutant déqualifié, suivant les prescriptions de la machine. Ces possibilités dépendent des conditions de déploiement et d'accompagnement de la technologie. Pour éviter l'aliénation et la prolétarisation, il ne faut pas mettre en place de dispositif automatique exerçant un contrôle complet sur l'humain (sinon, autant l'automatiser jusqu'au bout) : l'humain doit garder la capacité de reprendre la main.

IMPACTS ÉCONOMIQUES
&
DISTRIBUTION DE VALEUR

- L'intégration de l'IA au sein de nos économies est porteuse de **gains de productivité importants** et permet de créer davantage de valeur.
- Dans des économies ouvertes comme les nôtres, il y a un un **fort enjeu de compétitivité** et donc une urgence à intégrer rapidement ces innovations.
- Les ruptures récentes de l'intelligence artificielle résultent de la conjugaison de trois facteurs : un accès à des données massives, une puissance de calcul importante et des algorithmes sophistiqués. La recherche et le développement des algorithmes de *machine learning* et de *deep learning* se font très largement sur un mode *open source* : les algorithmes sont disponibles. **L'enjeu se situe dans l'accès aux données et à la puissance de calcul**.
- **Risque de dépendance sur le long terme et de perte de souveraineté** : les entreprises françaises et européennes pourraient à terme n'être que des consommatrices de logiciels et de solutions développées à l'extérieur.
- Dans ce contexte de recherche ouverte, **l'accès aux données nécessaires aux algorithmes devient donc un avantage compétitif**. Les plateformes disposent d'une position favorable. Risque de concentration de la valeur.
- L'IA peut enfin conduire à renforcer les **inégalités territoriales** – prime aux métropoles – qui reconduisent des logiques plus générales propres à l'économie numérique, concentration de la production dans les centres urbains, etc.

L'intégration de l'intelligence artificielle au sein de nos économies est vectrice de gains de productivité importants. En tant qu'innovation technique, elle constitue un intrant pour les entreprises à la fois dans leurs processus internes (management, logistique, service client, assistant, etc.) et dans leurs productions que ce soit de biens de consommation (objets intelligents, véhicules autonomes, etc.) ou de services (banque, assurance, juridique, médical, etc.). Ne pas intégrer ces technologies représenterait un risque de perte de compétitivité majeur.

Trois risques paraissent néanmoins devoir être pris en compte, en termes de dépendance économique, de souveraineté et de possible renforcement des inégalités. Dans la configuration actuelle de l'économie numérique, le risque que les entreprises françaises ne soient que des consommatrices de solutions développées à l'étranger est important. Le développement de l'intelligence artificielle pourrait alors

renforcer largement les effets de plateformisation et de captation de la valeur qui sont déjà caractéristiques de l'économie numérique : en effet, ce développement nécessite des moyens, en termes de données et de capacités de calcul, dont peu d'acteurs disposent. De plus, ces technologies pouvant servir d'outils d'aide à la décision, voire de prise de décision, il est nécessaire de préserver la capacité à en déterminer le sens et le contenu. Enfin, la diffusion de ces technologies peut avoir un effet important sur l'augmentation des inégalités territoriales et sociales, du fait notamment de la concentration de la valeur dans certains lieux.

Enjeu de compétitivité internationale

L'intelligence artificielle pourra apporter des gains de productivité (*cf.* annexe 3 "L'intelligence artificielle et l'enseignement des technologies génériques") et un surcroît de valeur à la production. Dans une économie ouverte, comme celle de la France, la production nationale se devra d'intégrer rapidement cette innovation pour rester compétitive autant en termes de prix (gains de productivité) que de qualité.

Comme d'autres innovations de l'ère numérique, l'intelligence artificielle est porteuse de changements radicaux dans les secteurs économiques. La France risque de se trouver confrontée à un conflit de transition, avec de grandes entreprises qui n'ont pas les moyens (stratégique, logistique, organisationnel) d'intégrer efficacement les innovations et qui utilisent des verrous institutionnels pour ralentir le développement de nouvelles entreprises plus innovantes. *In fine,* ce sont les entreprises qui ont pu se développer dans un cadre propice qui, en entrant sur le marché français, apportent les innovations et la disruption. La transition peut alors être assez brutale avec une conquête massive de la part des clients grâce à des démarches dématérialisées peu contraignantes, comme ce fut le cas avec Uber.

Recommandation : inciter les entreprises, grandes ou petites, à développer et intégrer des briques d'intelligence artificielle, notamment en fédérant les acteurs des différents écosystèmes autour d'un réseau thématique ; celui-ci mettrait à disposition, de façon contrôlée, les données nécessaires et faciliterait l'accès à l'expérimentation (comme pour le véhicule autonome).

Des réseaux de neurones en *open source*

En juin 2016, le magazine Wired prophétisait « *la fin du code* » en expliquant qu'à l'avenir, nous ne programmerons plus des ordinateurs : nous les « *dresserons* » à la manière dont nous pouvons dresser un chien[9]. De ce point de vue, l'apprentissage machine (*machine learning*) – en particulier les techniques d'apprentissage profond, basées sur des réseaux de neurones – a donné des résultats remarquables, en rupture majeure par rapport à l'existant, au cours des dernières années. Pour y parvenir, les algorithmes d'apprentissage profond nécessitent d'allier des moyens de calcul puissants à un accès à des données massives (*big data*).

La recherche et le développement des algorithmes de *machine learning* et de *deep learning* se font très largement sur un mode *open source*. Sur Github, une plateforme de mise à disposition de programmes, les

[9] Wired, juin 2016 : https://www.wired.com/2016/05/the-end-of-code/

projets en apprentissage machine figurent parmi les plus actifs[10]. Parmi ces projets, *TensorFlow*, qui regroupe les projets logiciels en apprentissage machine des équipes de Google, est rapidement devenu le répertoire le plus populaire sur la plateforme[11]. La France n'est pas en reste puisque la bibliothèque *Scikit-learn*[12], regroupant des outils d'IA développés en *open source* et également très utilisés, y a été largement développée. De la même manière, Facebook, Microsoft, IBM et Amazon ont opté pour un modèle *open source*, mettant ainsi gratuitement leurs travaux à la disposition de tous. Même Apple, d'ordinaire réputée pour son extrême discrétion, a annoncé en décembre dernier l'ouverture de ses recherches en intelligence artificielle[13]. Ces entreprises partagent régulièrement l'avancée de leur recherche avec la communauté scientifique. Leurs chercheurs[14] s'expriment également lors de conférences. À l'image de leurs homologues américains, les géants chinois du numérique ont eux aussi entrepris d'ouvrir leurs algorithmes d'intelligence artificielle : c'est le cas notamment, pour Baidu[15] qui a mis à disposition ses outils d'apprentissage profond en *open source*.

Risques de dépendance et de plateformisation

L'ouverture des algorithmes permet à la communauté des développeurs de progresser sur une technologie complexe qui demande une mobilisation collective. Néanmoins, seuls certains acteurs (Google, Amazon, IBM, etc.), disposant d'une capacité de calcul suffisante, sont en mesure de proposer aux entreprises des services d'intelligence artificielle – notamment hébergées dans le cloud.

Watson, l'intelligence artificielle développée par IBM, est déployée dans certaines entreprises pour assister les collaborateurs qui peuvent lui poser des questions en langage naturelle. Dans le cabinet d'avocat Baker & Hostetler, Ross, la version spécialisée en droit de Watson, est utilisée par l'équipe chargée des dossiers de faillite. Dans un autre domaine, Alexa, l'assistant personnel d'Amazon, qui se place au cœur des foyers pourrait devenir une porte d'entrée vers les clients incontournables pour de nombreux services et commerces, qui deviendraient ainsi complètement dépendant de cet intermédiaire.

Ces entreprises ont donc la capacité de se placer au cœur d'une activité et en interface entre une entreprise et ses clients. Elles peuvent ainsi reproduire le modèle des plateformes numériques comme Booking, Airbnb ou Uber qui ont su capter une part significative de la valeur d'un marché en se positionnant comme intermédiaire incontournable. Ce risque de plateformisation, typique du développement de l'économie numérique, induit donc des risques majeurs en termes de dépendance des acteurs économiques. L'avènement de l'intelligence artificielle déboucherait ainsi sur une concentration encore plus importante de la valeur économique – qui s'accompagnerait de conséquences sociales négatives (par

---

[10] https://github.com/showcases/machine-learning

[11] http://donnemartin.com/viz/pages/2015

[12] http://scikit-learn.org

[13] http://www.theverge.com/2016/12/6/13858354/apple-publishing-ai-research-siri-self-driving-cars

[14] Cette large ouverture de la recherche mondiale mérite néanmoins d'être relativisée : il semble qu'à l'inverse de la recherche occidentale, la recherche chinoise — aujourd'hui la plus active au monde ! — forme une communauté plus distincte. Des barrières linguistiques découle en effet une forme d'asymétrie : les chercheurs chinois, qui parlent généralement anglais, bénéficient d'un accès aux travaux rédigés dans cette langue. À l'inverse, la communauté anglophone accède plus difficilement aux travaux issus de la recherche sinophone.

[15] http://www.theverge.com/2016/9/1/12725804/baidu-machine-learning-open-source-paddle

exemple, moindres recettes fiscales). La dépendance de l'économie à des solutions d'intelligence artificielle développées à l'étranger peut également conduire à une perte de souveraineté, dès lors que les règles régissant les avancées technologiques échappent largement au contrôle national. Une fois les innovations établies et déployées à l'étranger, il devient très difficile d'en infléchir le fonctionnement ou de déterminer d'autres règles pour les régir.

Impact industriel de l'intelligence artificielle :
exemple du véhicule autonome

La médiatisation à partir de 2010 de projets de véhicule autonome portés par des entreprises du numérique comme Google, Apple, Tesla ou encore Uber a mis en lumière les avancées technologiques liées à l'intelligence artificielle dans le secteur de la production automobile. L'horizon du déploiement de véhicule autonome se rapproche avec des premiers usages non expérimentaux prévus dès 2020 et une diffusion à grande échelle envisagée à partir de 2030.

L'équipement des véhicules en intelligence artificielle de conduite autonome pose la question du partage de la valeur avec les constructeurs automobiles. De 40 % de la valeur ajoutée d'un véhicule, l'équipement logiciel pourrait passer à 70 % d'ici dix ans. Les constructeurs automobiles qui comptent sur leur savoir-faire en termes de design et de moteur pour maintenir leur image de marque sauront-ils développer des intelligences artificielles suffisamment performantes ou devront-ils intégrer des logiciels extérieurs qui pourraient se placer en nouvel intermédiaire avec leurs clients ? Au-delà de leur capacité à développer une intelligence artificielle, les consommateurs seront-ils plus enclins à acheter un véhicule équipé d'un logiciel du constructeur ou de celui d'une entreprise numérique ?

En fonction du rapport de force qui s'installera, c'est-à-dire de la marque qui s'imposera comme celle du véhicule, le partage de la valeur se fera en faveur soit du constructeur automobile soit du fournisseur d'intelligence artificielle.

Au-delà de cette bataille dans les véhicules, c'est à une remise en cause de toute l'organisation économique des transports routiers à laquelle il faut se préparer : impact sur les assurances, le permis de conduire, les garagistes, les sociétés de transport, la transition potentielle d'un système où les ménages sont propriétaires de leur véhicule vers un système de "location de véhicule à la carte", complémentarité entre les transports en commun actuels et les véhicules autonomes[16]

L'accès aux données devient un avantage comparatif

Dans le contexte d'une recherche et d'un développement de l'intelligence artificielle largement ouverts, la question posée est moins celle de l'accès aux algorithmes que celle de l'accès aux données pour les

[16] Janin L., Nemri M. et Raynard C., « La voiture sans chauffeur, bientôt une réalité », *La Note d'analyse* n°47, France stratégie.

entraîner. Celles-ci sont alors un élément crucial du développement de ces machines auto-apprenantes : sans une masse suffisante de données, les algorithmes d'apprentissage profond ne pourraient pas atteindre des taux d'erreur satisfaisants. En matière de *deep learning*, plus une machine est entraînée, plus elle est pertinente. Un algorithme de médiocre qualité pourrait ainsi être supérieur à un bon algorithme s'il est mieux entraîné.

Au sein des organisations, entreprises et administrations, les enjeux et l'importance des mégadonnées (*big data*) semblent de plus en plus compris et assimilés. Ce qui l'est beaucoup moins, c'est l'importance cruciale des données brutes en matière d'intelligence artificielle. Prises en grand nombre, ces données, même non raffinées, peuvent servir à entraîner des machines à détecter des corrélations pour, à moyen et long termes, prendre des décisions sur la base de celles-ci. De fait, les organisations qui ont accès à de larges bases de données sont en possession d'un actif crucial.

Recommandation : sensibiliser les organisations à la valeur des données en tant que ressources essentielles à l'entraînement des algorithmes d'intelligence artificielle.

Il est indéniable que les géants du numérique – qu'ils soient chinois, russes ou américains – dont le succès s'explique largement par leur capacité à collecter les données, notamment grâce à une taille de marché importante, partent avec un certain avantage dans la course mondiale à l'intelligence artificielle. Cela peut accroître les risques de dépendance économique et les inégalités en termes de répartition du pouvoir, de manière générale, et de la valeur économique, en particulier. Les géants du numérique ont alors tout intérêt à ouvrir les algorithmes de *machine learning*, pour capitaliser sur les logiques de l'innovation ouverte, tout en gardant la main sur les actifs qui en constituent la clé. Si ces grandes entreprises nouent aujourd'hui de nombreux partenariats avec des organismes de recherche pour alimenter en données les algorithmes et les améliorer, elles se gardent bien d'ouvrir toutes leurs données. Elles conservent ainsi la main sur leur actif stratégique dont elles déterminent les conditions d'accès éventuel et les modalités d'utilisation.

Conséquences du développement de l'IA sur le marché de l'externalisation

L'intégration de l'IA dans les processus de production conduit à modifier les chaînes de valeur. Différents effets peuvent être relevés. À l'échelle internationale, il est possible que des processus qui avaient été externalisés vers des pays à bas coût de main-d'œuvre soient à nouveau internalisés, dès qu'ils peuvent être réalisés par une IA. Ce processus ne met pas en danger l'emploi dans les pays développés, puisqu'il était déjà externalisé. En revanche, il peut constituer une menace pour certains pays qui se fondent uniquement sur une main-d'œuvre peu qualifiée spécialisée dans des tâches cognitives répétitives. Ce processus est clairement à l'œuvre pour certaines activités bancaires ou dans les centres d'appel.

Un autre processus est également à l'œuvre avec le développement des plateformes de micro-tâches, telles qu'Amazon Mechanical Turk ou FouleFactory. Sur ces plateformes, les développeurs d'intelligence artificielle ont recours à des travailleurs pour entraîner les algorithmes, *via* des tâches de

labellisation, de nettoyage de bases de données, de reconnaissance d'images…

La France et l'Europe doivent prendre conscience de leurs atouts

La question de l'accès aux données risque de se faire encore plus pressante à mesure de l'émergence d'une intelligence artificielle de moins en moins cantonnée à des tâches spécifiques (ce que l'on désigne par intelligence artificielle faible), i.e. une intelligence artificielle capable d'apprendre de nouvelles tâches ou d'acquérir de nouvelles capacités par transfert, voire par « capillarité ». Par exemple, DQN, un algorithme développé par DeepMind (Google) parvient à comprendre par lui-même les règles de différents jeux auxquels il n'a jamais été entraîné à partir de son entraînement sur un jeu spécifique. Si l'émergence d'une véritable intelligence semi-forte[17] n'est pas attendue avant plusieurs décennies, les capacités de l'intelligence artificielle actuelle tendent donc bien à s'horizontaliser et, de ce point de vue, les grandes plateformes bénéficient d'une position favorable. La diversification de leur activité, le dynamisme des acquisitions et des partenariats que les plateformes concluent les conduisent à amasser des informations dans de nombreux secteurs, ce qui est propice à l'horizontalité, c'est-à-dire à la capacité d'effectuer différentes tâches. La richesse d'une donnée tient en effet pour beaucoup à son potentiel de croisement avec d'autres.

Dans cette compétition mondiale, en France et en Europe les atouts ne manquent pas. En France, par exemple, les bases médico-administratives de santé sont alimentées chaque année par 1,2 milliard de feuilles de soin, 500 millions d'actes médicaux et 11 millions d'hospitalisation. D'autres données, comme les données juridiques, pourraient gagner à être ouvertes de manière plus large.

Recommandation : poursuivre et accentuer l'effort en matière d'accès aux données publiques et para-publiques.

Néanmoins, il ne s'agit pas de se focaliser uniquement sur les données publiques ou para-publiques, mais également, à plus long terme, d'engager une réflexion sur les modes de partage de la valeur entre acteurs privés. En ce sens, le CNNum s'est saisi de la question de la libre circulation des données aux niveaux international et européen, d'un point de vue de transfert entre les territoires mais également entre acteurs économiques.

Recommandation : approfondir la question de la libre circulation des données, entre espaces géographiques mais également entre acteurs économiques et en définir les conditions.

Inégalités territoriales au niveau national

[17] On distingue l'intelligence artificielle faible, capable de reproduire une tâche spécifique (celle qui est déployée aujourd'hui), l'IA semi-forte, capable de réaliser des tâches plus générales en comprenant notamment le contexte (un sujet de recherche) et l'IA forte, qui égale, voire dépasse, l'intelligence humaine (et relève toujours, à ce jour, de la science-fiction).

À l'échelle du territoire national, l'IA pose les mêmes questions de prime aux métropoles et d'inégalités territoriales que celles posées par le développement économique des dernières années[18]. La numérisation fait courir le risque que les activités de production se concentrent dans des centres urbains de taille suffisante pour bénéficier des externalités de réseau, notamment pour la production d'IA car celle-ci suppose de réunir des facteurs de production localisés principalement dans les métropoles (capital humain, R & D, capital physique, écosystème). En revanche, en ce qui concerne l'utilisation des services permis par une IA, les gains escomptés peuvent être répartis sur tout le territoire. Pour ce faire, ses services doivent être diffusés dans les territoires, en veillant à ce qu'ils soient accessibles au plus grand nombre.

[18] Lainé F. (2017), « Dynamique de l'emploi et des métiers : quelle fracture territoriale ? », *La Note d'analyse* n° 53, février, France stratégie.

PENSER LA COMPLÉMENTARITÉ
Humain-Machine

- Réfléchir en termes de tâches plutôt que de métiers permet de **se pencher sur la transformation plutôt que sur la destruction d'emplois**.
- Afin d'appréhender ces transformations, il est urgent de définir la gouvernance permettant d'anticiper ces impacts, notamment sur trois points clefs :
    - mesurer la destruction comme la transformation des emplois, et donc **évaluer les compétences substituables au regard d'un faisceau de critères** :
        - *La technologie est-elle suffisamment avancée pour que cette tâche soit automatisée ?*
        - *La tâche nécessite-t-elle des capacités cognitives verticales (orientées sur une tâche très spécifique) ou horizontales ?*
        - *L'automatisation de cette tâche est-elle acceptable socialement ?*
        - *Cette tâche requiert-elle le recours à une intelligence émotionnelle ?*
        - *Cette tâche nécessite-t-elle une intervention manuelle complexe ?*
    - Déterminer les compétences qui demeurent à l'avantage de l'humain peut ne pas suffire, il est nécessaire de **réfléchir à la complémentarité humain-machine** au niveau de l'individu et de l'organisation.
    - À partir de la détermination des compétences non substituables, **une réforme en profondeur de la formation tout au long de la vie** doit être effectuée, en termes de temporalité, de contenu et de structures.

Le développement de l'IA aura également un impact sur les conditions de travail des individus et la structure des organisations. Les études macro-économiques se penchent surtout sur la question de la substituabilité de la machine par rapport au travail humain. Il serait évidemment hors de propos de balayer d'un revers de main l'hypothèse d'une diminution irréversible du nombre d'emplois. Pour autant, et même si cela n'obère pas des réflexions parallèles sur le mode de redistribution de la valeur, il nous semble nécessaire de réfléchir à la transformation des métiers plutôt que d'acter cette diminution irréversible. En effet, cette hypothèse raisonne à tâches constantes. Or, comme nous l'apprend le rapport du COE qui prend une approche par tâches et non par métiers, un certain nombre de tâches vont être automatisées tandis que d'autres ne le seront pas, celles où réside l'avantage comparatif de l'être humain. En effet, selon le COE, 50 % des emplois verront la moitié des tâches qui les composent automatisés, et 10 % des

emplois verront l'intégralité de leurs tâches menacées. Cela signifie que l'on peut partir du constat que les métiers seront en grande partie transformés plutôt que détruits. Ce constat peut s'appuyer sur les enseignements issus des précédentes révolutions industrielles, qui ont vu – parfois non sans une transition difficile – de nouveaux emplois apparaître en remplacement de ceux qui étaient détruits. De plus, l'efficacité d'un algorithme associé à un humain est bien plus importante que celle d'un algorithme seul : l'IA prend en charge les tâches répétitives, et l'humain peut se concentrer sur les tâches les plus riches. Pour ces deux raisons, il semble nécessaire de réfléchir à la notion de complémentarité homme-machine. Comment s'organiser pour mener ces réflexions de manière collective ?

Organiser une gouvernance en vue d'anticiper ces transformations

La réflexion sur les impacts économiques et sociaux de l'AI ne peut se faire de manière descendante au regard de l'ampleur des transformations qu'elle est susceptible d'engager. Il s'agit, à tous les niveaux, de définir la gouvernance qui structure un dialogue et des prises de décision nécessaires pour anticiper les impacts de l'IA, et de réfléchir à la répartition des gains de productivité permis par le numérique.

À titre d'exemple, cinq échelons pourraient être dans un premier temps ciblés : le niveau national, le niveau régional, le bassin d'emploi, la branche, l'entreprise. À chacun de ces niveaux, un travail préalable d'identification du réseau des acteurs pertinents serait mené, pour organiser une concertation et la construction participative de projets pertinents.

Une plateforme (numérique et sous forme de concertation nationale) avec tous les acteurs de l'emploi, Pôle emploi, AFPA, acteurs régionaux, nationaux, européens, syndicats, serait mise en place à cet effet.

Trois sujets prioritaires feraient l'objet de cette concertation nationale :
- réfléchir à la complémentarité entre les compétences automatisables et celles non automatisables, afin de déterminer les compétences humaines substituables et de maîtriser la transition ; .
- de manière corrélative, identifier des modes de complémentarité vertueux qui puissent être pérennes et qui ne soient pas créateurs d'aliénation supplémentaire au niveau individuel et collectif ;
- enfin, repenser de fond en comble les modes de formation tout au long de la vie, du fait de la complémentarité évoquée ci-dessus.

Quelles compétences substituables ?

De nombreuses études se sont attachées à définir les compétences humaines qui, à ce jour n'apparaissent pas substituables par l'IA (cf. annexe 1 "L'intelligence artificielle en quête d'acceptabilité et de confiance"). Les compétences englobant une dimension affective et relationnelle, certaines compétences physiques de précision et surtout, les capacités d'analyse générale (flexibilité, capacité d'adaptation, capacité à résoudre les problèmes, créativité), semblent *a priori* demeurer accomplies de manière plus efficace par l'homme.

Plus précisément, il est urgent de mesurer la substituabilité des tâches au regard d'un faisceau de critères, techniques *et* sociaux, et prenant en compte un contexte économique allant au-delà de la seule organisation de l'entreprise ou du secteur d'activité :

- La technologie est-elle suffisamment avancée pour que cette tâche soit automatisée ?
- La tâche nécessite-t-elle des capacités cognitives verticales (orientées sur une tâche très spécifique) ou horizontales[19] ?
- L'automatisation de cette tâche est-elle acceptable socialement ?
- Cette tâche s'appuie-t-elle le recours à une intelligence émotionnelle ?
- Cette tâche requiert-elle une intervention manuelle complexe ?

Les compétences non substituables doivent être développées (*cf.* partie formation). Néanmoins, ce développement, s'il est nécessaire, n'est pas suffisant pour assurer une complémentarité vertueuse entre l'homme et l'intelligence artificielle. De nombreux facteurs influent sur les modalités de cette complémentarité, au premier titre desquels l'organisation du travail. L'introduction de l'intelligence artificielle peut avoir des impacts négatifs au niveau individuel et collectif. Certains risques sont décrits dans l'annexe 1 "L'intelligence artificielle en quête d'acceptabilité et de confiance".

Penser une bonne complémentarité humain-machine

Définir des principes de complémentarité permettant de garantir l'acceptabilité et l'appropriation de l'IA, une diffusion pérenne de ses apports, une protection contre les risques systémiques qu'elle peut contenir en germe, s'impose aussi.

Par ailleurs, au niveau individuel, le travail quotidien avec l'IA au-delà de ses nombreux atouts, peut également présenter des risques sociaux et psychiques. La concentration des capacités d'analyse et d'intelligence dans les nouveaux outils d'IA renferme le risque de mener à une "prolétarisation"[20] d'un nombre encore plus grand de travailleurs, au sens de la privation de leurs savoirs qui valorisent leurs compétences. La délégation d'un nombre toujours plus important de tâches à l'IA est également susceptible d'engendrer un désengagement, une déresponsabilisation des travailleurs, du fait de l'appauvrissement des relations interpersonnelles ainsi que d'une délégation accrue de la prise de décision.

Au niveau des organisations, la diffusion des technologies liées à l'IA devra être appréhendée en termes de risques cognitifs, notamment d'accroissement potentiel de la propension à la "stupidité fonctionnelle". Une introduction massive de l'IA dans les organisations peut favoriser l'obéissance, la conformité et la prévisibilité des comportements. En effet, si la diffusion de l'IA entraîne l'automatisation d'un nombre encore plus élevé de tâches et d'activités répétitives, elle peut dans le même temps contribuer à la délégation de la prise de décision et favoriser des contextes organisationnels diminuant la mobilisation des capacités cognitives. Elle pourrait ainsi favoriser le manque de réflexivité des individus et de l'organisation ainsi que l'absence de justification des décisions. En optimisant toutes les tâches, elle réduirait, voire annihilerait les capacités à innover. De telles situations mèneraient alors au développement

[19] Cette distinction verticale/horizontale renvoie à la distinction entre IA faible et IA forte, développée p. 9

[20] Pour reprendre la notion développée par Bernard Stiegler dans son ouvrage *La société automatique*, Fayard, 2015.

de risques systémiques, qui menaceraient, au-delà de l'organisation, les structures économiques et sociales.

Encadré cas pratique :
l'exemple du centre d'appel

Le cas du métier d'opérateur de centre d'appel est intéressant car il montre l'ambivalence de l'automatisation au sein des organisations :

- **Scénario où le travailleur est « augmenté » par l'IA :**

  - Automatisation des appels les plus simples : 80 % des appels sont automatisés.
  - Concentration de l'opérateur sur les 20 % d'appels restants, les plus complexes et qui permettent de développer des compétences commerciales et relationnelles.
  - Sur les 20 % d'appels restants, l'opérateur est aidé par l'IA : il peut utiliser des moteurs de recherche sémantiques permettant de retrouver très rapidement l'information, et, à nouveau, de se concentrer sur la relation client.

- **Scénario où le travailleur est subordonné à l'IA :**

  - Une intelligence artificielle analyse les appels clients et propose un script d'interactions à l'opérateur, qui se contente de lire à haute voix les réponses proposées par la machine. Ces réponses sont optimisées au regard de l'analyse de l'ensemble des autres interactions passées.
  - L'opérateur devient donc un support vocal et émotionnel en réponse quasi-automatique à l'intelligence artificielle.

Se former dans un monde d'intelligence artificielle

Au regard de l'objectif d'une complémentarité vertueuse entre l'IA et le travailleur, la formation joue un rôle incontournable : en effet, développer des compétences transversales de créativité, d'adaptabilité et de capacité à résoudre des problèmes rend nécessaire une réforme importante de la formation tout au long de la vie, même si celle-ci n'est pas uniquement liée aux modifications du marché du travail induites par l'intelligence artificielle. La formation professionnelle fait régulièrement l'objet de nombre de critiques concernant ses modes de fonctionnement et son efficacité. De plus, l'exigence d'un décloisonnement entre formation initiale et continue, au regard de la difficulté d'anticiper les activités et les métiers de demain, a déjà été défendue dans un certain nombre de travaux, notamment ceux du CNNum sur la transformation numérique de l'enseignement supérieur et la recherche[21]. Néanmoins, les bouleversements massifs que va connaître le marché du travail du fait du développement de l'IA pose la question de la formation tout au long de la vie avec une acuité nouvelle. **En effet, en accélérant l'obsolescence des compétences acquises lors de la formation initiale, l'intelligence artificielle pose la question de l'accroissement du temps de formation continue et de la modification de ses contenus et de ses structures.** Le lien entre le fait d'avoir suivi une formation initiale et celui d'exercer le métier correspondant toute sa vie est

[21] Cf. travaux du CNNum sur la transformation numérique de l'enseignement supérieur et de la recherche : Référentiel de transformation numérique de l'Enseignement supérieur et de la Recherche, 19/12/2016.

clairement à remettre en cause : d'où une formation tout au long de la vie qui puisse véritablement accompagner les transitions, à l'intérieur d'un même métier mais également vers les métiers et les secteurs d'avenir.

Or le ratio actuel entre temps de travail et temps de formation pourrait ne pas être à la hauteur de l'exigence de renouvellement constant des compétences. C'est pourquoi il faut réfléchir à l'extension des temps de formation continue, pour donner à chacun la possibilité d'acquérir les nouvelles capacités techniques et non techniques dont auront besoin les individus dans un monde d'intelligence artificielle. Cette extension des temps de formation continue devra aller de pair avec la mise en place d'une continuité plus importante entre formation initiale et formation continue. **Les professions potentiellement les plus touchées par l'automatisation pourraient être la cible prioritaire d'une politique de formation préventive.**

La modification des contenus de formation s'impose ensuite. Le défi de la formation initiale et continue consiste à trouver l'équilibre entre l'acquisition de savoirs et compétences en phase avec les besoins actuels de la société et la nécessité de préparer un avenir dont on sait au moins qu'il exige de tous un esprit critique, une capacité d'apprendre à apprendre. Le développement de l'intelligence artificielle renforce cet état de fait. Il nécessite à la fois de permettre l'acquisition des compétences techniques qui lui sont spécifiques afin que les besoins de main-d'œuvre du secteur puissent être remplis, mais également de développer les qualités générales d'adaptabilité, de créativité et d'esprit critique. À terme, ces compétences peuvent seules permettre aux individus de développer leur complémentarité avec la machine. Cela signifie qu'outre les formations spécifiques à l'intelligence artificielle, la question de la réorientation de la formation continue vers des formations généralistes, permettant de développer des compétences transversales d'adaptabilité et de créativité, devra être posée. Le développement de ces compétences nécessite de réinventer les modes de formation (formation en contexte, sur poste, en conception) et le rapport au savoir lui-même, autour de pratiques de partage, de mise en réseau, d'expérimentation, de projets…

La transformation des structures de la formation tout au long de la vie est enfin inévitable. Au regard de la nécessité accrue, pour chaque travailleur, de pouvoir se former selon ces exigences, les entités existantes ne sont ni suffisantes ni adaptées. Le rôle de l'Université dans la délivrance de la formation tout au long de la vie pourrait être ainsi davantage affirmé, pour assurer à tous une formation de qualité. Pôle Emploi évoluerait pour s'inscrire dans une logique préventive. Enfin les dispositifs de formation informels (tiers-lieux, fablabs, makerspaces…) et d'expérimentation de pratiques pédagogiques nouvelles doivent être intégrés pleinement à l'évolution des structures de formation tout au long de la vie.

Recommandation : viser une transformation d'envergure de la formation tout au long de la vie, en termes de ratio temps de travail/temps de formation, de contenus et de modes de formation ainsi que de structures porteuses.

Annexe 1

L'intelligence artificielle
en quête d'acceptabilité et de confiance

Contribution au groupe de travail : Anticiper les impacts économiques et sociaux de l'intelligence artificielle

Laurence DEVILLERS (Sorbonne/CNRS, CERNA)
Danièle BOURCIER (CNRS, CERNA)
Stéphane BUFFAT (médecin, IRBA)
Alexei GRINBAUM (CEA, CERNA)
Arnaud LAVIT D'HAUTEFORT (cabinet Ytae)
Christophe LAZARO (Université de Namur, Belgique, CERNA)

De manière générale, il faut souligner le caractère ambivalent de toute technologie qui offre de nouvelles opportunités en même temps qu'elle crée des risques. Il faut ne pas négliger les aspects éthiques du développement de ces technologies. Les systèmes d'intelligence artificielle deviennent aujourd'hui capables d'accomplir certaines activités qui auparavant étaient l'apanage exclusif des humains. Le développement de ces systèmes conduit au renforcement progressif de leurs facultés d'autonomie propre et de cognition – la capacité à apprendre par l'expérience et à prendre des décisions de manière indépendante –, lesquelles sont susceptibles de faire de ces systèmes des agents à part entière pouvant interagir avec leurs opérateurs et leur environnement et les influencer de manière significative. Ce développement technologique est de nature à transformer considérablement la sphère de l'emploi et du travail, au point d'engendrer une nouvelle phase de révolution informatique dans l'industrie et dans les services. Celle-ci aura notamment pour conséquence la multiplication des systèmes apprenants dans l'environnement de travail, l'intensification des collaborations entre humains et machines ainsi que la délégation croissante de tâches et d'activités autrefois réalisées par les humains à des machines. Ces collaborations entre humains et machines vont revêtir toutes sortes de formes et seront différentes des collaborations classiques.

Cette conjoncture exige qu'on étudie les effets potentiels de l'IA en termes d'éthique et d'acceptabilité sociale. À ce titre, il est nécessaire de soulever une série de questions et de s'interroger sur la formation à l'IA, la transformation du travail grâce à l'IA, l'acceptabilité des mutations et la régulation de l'IA. Voici les questions que la société doit se poser avant de décider du pourcentage d'IA susceptible d'accroître la rentabilité pour que chacun reste le plus possible utile à la société par son travail ou ses compétences :

- Comment l'IA va-t-elle modifier la conception de l'emploi et du travail ?
- Que faut-il préserver des valeurs du travail humain ?
- Les humains vont-ils accepter de travailler avec des IA ou sous les ordres de systèmes d'IA ?

- Serons-nous toujours ceux qui supervisent/surveillent ?
- Comment évaluer la responsabilité d'un agent humain ou numérique dans un monde où l'humain n'aura désormais plus le monopole de la décision et de l'action ?
- Comment créer les conditions d'une coordination des activités et d'une collaboration saines entre l'homme et les systèmes d'IA ?

Cette annexe à la note *Anticiper les impacts économiques et sociaux de l'intelligence artificielle* entend proposer un certain nombre de pistes de réflexion destinées à construire à court et moyen terme une relation de confiance entre les travailleurs, les organisations et les systèmes d'IA. Elle a également pour objectif de formuler diverses recommandations visant à garantir le respect des valeurs humaines aux stades de la conception, de l'acceptation, de l'appropriation et de l'utilisation de ces systèmes.

I. Champ d'application

a) Les métiers bouleversés par l'IA

L'IA va bouleverser de nombreux métiers, comme l'ont fait les machines-outils lors des révolutions industrielles précédentes. D'innombrables tâches, actions et fonctions pourront être réalisées et supervisées par une IA. Des agents autonomes participent déjà à différents types de négociations, de transactions (exemple : transactions boursières), d'interactions avec les clients, les usagers, les entreprises et organisations. À court terme, vont être affectés de manière encore plus large les métiers du calcul et ceux qui suivent une procédure formellement établie. Certains métiers pourront être entièrement remplacés par des machines autonomes (exemple : conducteurs) tout en créant des nouvelles tâches de supervision. D'autres resteront entièrement le propre de l'homme car l'IA sera considérée comme inacceptable, inefficace ou non bénéfique pour la société. Dans d'autres cas encore, des formes inédites de coopération homme-machine verront le jour dans des environnements de travail où les IA seront des partenaires, des assistants plutôt que des remplaçants. De nouveaux métiers vont également apparaître comme ceux de concepteur, d'entraîneur ou de gestionnaire (cogniticien) de données et d'informations, surveillant ou encore réparateur d'IA, etc. **La valeur du travail ne sera plus le faire mais le contrôle, la vérification, l'approbation et la validation.**

Exemple : *On peut imaginer des phénomènes de retour à des métiers/tâches avec une réalisation partiellement manuelle et des résultats apparents. Vincent de Gaulejac : « Cette tendance peut s'expliquer par deux éléments. C'est un double changement, à la fois vers le manuel : du concret vers l'abstrait et il y a cette volonté de posséder à nouveau ce que l'on fait. Les cadres ont de plus en plus l'impression d'être pris dans un univers où leur travail ne fait plus sens. »* Paru dans *Le Monde.*

La rémunération reviendra aux concepteurs d'un procédé ou aux entraîneurs, aux vérificateurs et non pas à celui qui l'exécute car c'est la machine qui exécutera. La signature et la responsabilité finales pourraient en revanche rester humaines, c'est un choix. **Les systèmes d'IA ne prendront pas la place des humains dans toutes leurs tâches et fonctions. Il sera également nécessaire de créer de nouveaux métiers d'entraides sociales où l'humain sera au centre.**

Cette annexe met l'accent sur les métiers de service et vise, en particulier, à appréhender l'impact des IA sur certains secteurs de l'économie qui semblent plus "protégés" que les autres, mais qui ne sont pas pour autant à l'abri de changements significatifs.

Ces secteurs concernent les métiers revêtant une dimension :
- ***affective*** (exemple : soins infirmiers, assistance sociale, psychothérapie) ;
- ***corporelle*** (exemple : médecine, physionomie, analyse kinésique) ;
- ***mimétique*** (exemple : activités sportives, éducation) ;
- ***téléologique*** (exemple : management, justice, prêtrise) ;
- ***créative*** (exemple : activités artistiques et de recherche).

Ces différentes dimensions peuvent bien sûr se retrouver au sein d'une même activité ou profession.

b) Les phases de transformation du travail

Il existe deux échelles temporelles de transformation du travail :

1. Une lente évolution des métiers dans le contexte et l'environnement existants : les choix des utilisateurs portent sur "l'ancien" ou le "nouveau", mais l'ancien reste une norme parfaitement acceptée et disponible.

2. Une intégration totale des fonctions par le nouvel environnement de la vie humaine (ville intelligente, maison intelligente, objets connectés). Les contextes "anciens", sans disparaître complètement, deviennent archaïques ou exceptionnels (exemple : la maison haussmannienne non adaptée aux exigences technologiques modernes, le cheval remplacé par l'automobile). L'environnement "nouveau" s'impose en vertu des choix par défaut faits par la majorité, des décisions du secteur marchand ou des gouvernements. Ces options seront de toute façon stimulées par l'offre du marché. Dans la deuxième phase, les métiers subissent un changement de rupture, tandis que, dans la première, les compétences évoluent graduellement. Le passage de la première étape à la seconde n'est pas automatique, mais dépend de la réaction de la société, des changements générationnels et du niveau d'éducation.

L'acceptabilité des technologies apparaît dans ce processus comme un va-et-vient permanent entre les innovateurs et les divers types d'utilisateurs des nouvelles technologies.

- **Notion d'acceptabilité :** acceptabilité sociale (avant la mise sur le marché, tient compte de la pluralité des usages et perceptions) : essayer à travers des méthodes qualitatives (par exemple des scénarios) de cerner les questions d'usage que se pose le public: focus group, feedback à l'ingénieur.

Remarque : *Quel choix va réellement s'offrir à l'individu : un choix social, d'éducation et de formation ?*

- **Notion d'appropriation :** sur le plan individuel, on peut définir l'appropriation comme un processus qui repose sur l'acquisition progressive d'une certaine maîtrise technique et cognitive

des objets techniques ainsi que sur l'intégration significative de ces objets dans la pratique quotidienne.

Remarque : *Il y a deux niveaux d'appropriation : celui de l'interface de médiation humain-IA et celui de la compréhension fine des mécanismes sous-jacents. Cela peut être intéressant en cas de scénario « d'accident » pour décider de la responsabilité.*

II. Impacts éthiques et sociaux

Nous abordons ici les concepts et les idées dont la signification est susceptible d'évoluer sous l'effet de l'IA (responsabilité, mérite et valeur, créativité), puis les effets positifs (moins d'oubli, moins de perte de temps, prise en charge de tâches très répétitives ou des routines procédurales sans intérêt, etc) ; enfin, les effets susceptibles d'être négatifs (manque d'exercices de la mémoire, etc) si la société ne met pas de garde-fous et ne régule pas l'IA.

a) Changements de concepts

1. Responsabilité

L'entraîneur est celui qui réalise l'apprentissage du système d'IA en lui fournissant des données. L'entraîneur peut être le concepteur de la machine mais aussi celui qui a construit le corpus de données ou les utilisateurs conscients ou non conscients du fait que la machine apprend au cours de son utilisation. **La responsabilité sera partagée entre le concepteur, l'entraîneur et l'utilisateur selon des modalités qui restent à établir. Cela nécessite de modifier le droit pour prendre en compte ce type de partage.**

2. Mérite et valeur

La valeur du travail ne sera plus le *faire* mais le contrôle, la vérification, l'approbation et la signature. La rémunération reviendra au **responsable** d'un procédé et pas à celui qui l'exécute (car la machine exécutera).

3. Créativité

Un système d'IA convenablement programmé peut créer de l'innovation par expérimentation ou par sérendipité, mais seul l'humain donne un sens à la réalité nouvelle. Il est important de distinguer les créations des machines et les créations humaines.

b) Risques

Les aspects positifs de l'IA sont nombreux à condition de respecter un certain nombre de valeurs : une mémoire élargie, moins de perte de concentration ou d'oubli d'information pertinente, un conseil expert plus efficace, moins subjectif ou arbitraire, mieux justifié, etc, **à condition que la machine soit neutre, loyale et non discriminante**. Nous ne développons pas ici tous ces points déjà suffisamment soulignés et discutés par les acteurs de l'IA mais les risques éthiques. À chaque fois, il faut réfléchir aux aspects positifs et plus négatifs des situations et trouver les conditions pour transformer ces risques en atouts.

1. Désengagement : l'utilisation de l'IA et de machines autonomes peut conduire à un *désengagement* de l'humain. Remplacer le facteur humain par de l'IA pourrait conduire à une déshumanisation des pratiques et un appauvrissement des interactions sociales, lesquels constituent très souvent la raison d'être de certains métiers. C'est pour cela qu'il s'agit non pas de remplacer l'humain pour toutes les tâches sociales (comme par exemple dans la santé) mais de trouver les garde-fous pour accompagner au mieux la collaboration humain-machine. N'oublions pas que nombre de "décisions" sont actuellement prises automatiquement sans que le destinataire en soit informé. L'IA peut restituer au moins les traces et les raisonnements activés.

Exemple : le secteur de la santé et, en particulier celui des soins apportés aux patients malades, aux personnes en situation de handicap et aux personnes âgées va, dans les prochaines années, être profondément transformé par le développement de technologies offrant une grande diversité d'applications allant de l'assistance, au contrôle, jusqu'à l'accompagnement quotidien. La forte dimension affective au cœur de la relation de soin va évoluer et, dès lors, va modifier à la fois le travail des aides-soignants et thérapeutes ainsi que la vie des patients. Sans contester l'utilité indéniable des IA et des agents autonomes (robot portant, robot soignant ou robot coach) pour faciliter et soulager le travail des prestataires de soins, il est important de souligner que le contact humain est l'un des aspects essentiels de la relation de soin.

2. Déresponsabilisation : dans la continuité du point précédent, l'utilisation de l'IA et de machines autonomes peut entraîner une *déresponsabilisation* du travailleur, susceptible de ne plus devenir qu'un simple relais de la décision de la machine. Dans les cas de plus en plus nombreux de décisions prises sur la base d'un traitement automatisé de données, l'être humain pourrait avoir tendance à s'en remettre exclusivement à la proposition de la machine, en évitant d'engager sa responsabilité. Dévier de la solution préconisée par la machine entraînerait une prise de risque trop grande, susceptible de lui faire encourir d'éventuelles sanctions en cas de problème.

Exemple : le banquier qui s'en remettrait uniquement à la décision suggérée par la machine sur la base d'un profil pour accorder ou refuser un prêt à un client. Il y a également un problème de représentation des échelles : dans le trading par exemple, et la question de la numérosité à travers l'IA semble intéressante à creuser.

3. Transformation du mérite du travail : l'utilisation de l'IA et de machines autonomes peut bouleverser radicalement l'évaluation des performances des travailleurs et l'*idée de mérite*. En transformant les modalités de réalisation des tâches, les technologies feront évoluer les évaluations du mérite individuel et donc de l'équité des bénéfices induits par l'effort et des récompenses dues en vertu des talents acquis par l'individu.

Exemple : comment évaluer les performances d'un joueur d'échecs, d'un chirurgien ou d'un trader assisté d'une IA ? Qui est à la source de la performance ? Est-ce la technologie, le talent propre de l'humain ou l'association des deux ? Leur récompense (que ce soit en termes de salaire, de réputation) est-elle juste ? Est-elle méritée ?

4. « Prolétarisation » des savoirs et des savoir-faire : L'utilisation de l'IA et de machines autonomes peut contribuer à une *« prolétarisation » des savoirs et des savoir-faire*. À force de ne plus être entièrement partie prenante dans l'élaboration et l'expression d'une décision dans l'environnement de travail, les humains risquent de se voir déposséder de leur expertise en terme de *know-how,* et de perdre un ensemble de capacités et de compétences, qui non seulement peuvent être utiles à la société, mais qui contribuent aussi à alimenter le respect de soi. Ce risque pourra survenir si l'humain n'est pas associé à l'entraînement et à l'évaluation des modèles et au choix des données pour améliorer leur performance. **Se pose la question d'un conservatoire des métiers, avant l'émergence massive des intelligences artificielles.**

Exemple : le douanier dans un aéroport auquel on ne demandera plus d'exercer ses compétences et son sens de la physionomie acquis au fil des années à travers une expérience sur le terrain et qui devra s'en remettre aux prédictions de la machine en matière de profils de dangerosité.

5. Autonomie et créativité des travailleurs : l'utilisation de l'IA et de machines autonomes peut altérer l'*autonomie et la créativité des travailleurs***.** Derrière les vertus facilitatrices de certains dispositifs, peut se déployer de manière sous-jacente une normativité, prenant la forme d'un « paternalisme technologique » aux multiples visages : ainsi, les IA alertent, recommandent, rappellent à l'ordre, bloquent, interdisent, voire influencent.

Exemple : pour des raisons légitimes de sécurité ou d'organisation, certains dispositifs pourraient empêcher les êtres humains d'adopter des comportements considérés comme « sub-optimaux », de tenter certaines expériences, voire de commettre des erreurs qui souvent sont à la source de nouveautés et de découvertes.

6. **Vie privée et intimité au travail :** l'utilisation de l'IA et de machines autonomes peut nuire à la *vie privée et à l'intimité* des travailleurs en imposant un contrôle et une surveillance disproportionnés de leurs activités. Sans contester les bénéfices que peuvent procurer les IA dans la réalisation de multiples tâches, leur utilisation risque d'être mise en œuvre de manière abusive afin d'évaluer la productivité et le rendement des travailleurs et de les soumettre à un contrôle strict.

Exemple : pour les avocats dans un cabinet d'avocats ou les journalistes dans un groupe de presse et pour de nombreux autres métiers.

7. Épanouissement dans le travail : l'utilisation d'IA et de machines autonomes peut nuire à l'*épanouissement* des travailleurs sur le lieu de travail en perturbant le rythme de travail, les routines et en battant en brèche les conditions qui permettent de se réaliser dans le travail.

Exemple : le degré d'appréciation de la dangerosité d'un conducteur par un policier n'aura plus de valeur face aux contrôles automatiques. L'interruption du processus de pensée par une sollicitation non hiérarchisée de stimulation venant de plusieurs agents peut contribuer à désorganiser, voir à annihiler la capacité à prendre des décisions (Task Switching).

8. Disparités sociales affaiblies ou aggravées : l'utilisation de l'IA et de machines autonomes peut conduire à *affaiblir certaines disparités sociales (*exemple : valeur du "capital culturel"*) ou au contraire les aggraver* (les personnes non initiées perçoivent encore moins l'utilité de l'acquisition de ces connaissances à l'apparence "superflue").

Exemple : la machine mémorise pour le travailleur, du coup celui-ci fait de moins en moins d'efforts pour mémoriser les informations nécessaires pour résoudre un problème, comprendre et surveiller ce que fait la machine.

9. Pratique des managers et des représentants du personnel : dans une société qui se numérise, les pratiques du digital tendent à redéfinir les modèles d'interaction entre les individus. En entreprise, le numérique contribue *à redéfinir la pratique des managers et des représentants du personnel***,** qui jouent un rôle majeur d'interface entre la direction et les salariés.

Exemple : DIGILABSOCIAL - Digital Management & Relations sociales
Bénédicte Tilloy, Secrétaire générale et DRH de SNCF Réseau et Jean-Paul Bouchet, Secrétaire général de l'Union confédérale des cadres de la CFDT ont pris l'initiative de lancer avec Sciences Po et le Cabinet Ytae, un laboratoire d'idées permettant de mieux dessiner les contours de ce que seront demain les fonctions de manager et de représentant du personnel dans une entreprise digitalisée et de proposer des pratiques à expérimenter. Le laboratoire associe, à parts égales, une cinquantaine de managers et de représentants du personnel issus d'entreprises appartenant à des secteurs d'activités différents.

10. La "normalisation" d'une nouvelle technologie (au sens de Kuhn, où elle cesse d'être vue comme révolutionnaire) passe toujours par la normalisation (= standardisation). Or avec l'IA (comme avec les nanos), nous avons une technologie transversale qui ne se limite pas à un secteur industriel donné. D'où le dilemme : soit on instaure des normes par secteur, soit on crée des "normes d'IA" pour tous les secteurs.

Exemple : la France a expérimenté la dernière approche avec les nanos et cela a très mal marché. En revanche, l'Europe fait de la normalisation par secteur avec les nanos et cela se présente plutôt comme un succès.

11. Transformation des capacités humaines : les facilités qu'apporte la machine, par exemple sur la mémorisation, vont ***transformer certaines des capacités humaines.*** Devons-nous anticiper ce besoin dans l'enseignement ? (en soulignant l'utilité d'exercice sollicitant la mémoire court terme et long terme (exemple : calcul mental).

Remarque : *Aujourd'hui, j'oublie sans conséquence car la machine (téléphone, ordinateur) me rappelle ce que j'ai oublié. De ce fait, je n'exerce pas ma mémoire, mais mon attention aux indices de rappel… la machine aide à se souvenir mais rend dépendant (https://lejournal.cnrs.fr/articles/le-numerique-nous-fait-il-perdre-la-memoire).*

III. Recommandations

Afin de répondre aux impacts éthiques et sociaux des systèmes issus de l'IA et de faciliter leur adéquation et leur appropriation par tous les acteurs du monde du travail, il est important de réfléchir dès maintenant aux différentes étapes et conditions pour instaurer un climat de confiance.

La confiance se décline à un triple niveau :
(1) **la *conception*** : certaines valeurs doivent être prises en compte dès la phase de conception technologique et intégrées dans celle-ci ("value by design") ;
(2) **l'*appropriation*** : il est nécessaire de faciliter l'apprentissage par l'humain et son acquisition progressive de la maîtrise technique et cognitive des dispositifs techniques ainsi que l'intégration significative de ces dispositifs dans la pratique quotidienne ;
(3) **l'*utilisation*** : il faut informer les utilisateurs de l'origine automatique ou non de la décision, anticiper les conséquences possibles de l'adoption de ces systèmes de l'IA par les êtres humains et cerner en quoi ces systèmes vont redéfinir les modalités concrètes du travail.

Afin de garantir une transition plus aisée de la recherche à la commercialisation, nous proposons au gouvernement de promouvoir des programmes de recherche qui intègrent des mécanismes de vérification des résultats à court et moyen terme afin de comprendre les opportunités et les risques réels qui sont associés au déploiement et à la dissémination de ces technologies. La formation à l'IA et aux mutations est essentielle. La conduite du changement doit être menée en bonne intelligence avec tous les acteurs : dirigeants de société et syndicats.

Des ressources suffisantes doivent être consacrées à la résolution des défis éthiques et sociaux soulevées par le développement et les applications de ces technologies.

De manière générale, l'approche développée doit être graduelle, pragmatique et prudente. Dans cette perspective, nous formulons diverses recommandations.

1. Code d'éthique : l'éducation des ingénieurs à la réflexion éthique, par toutes les méthodes pédagogiques y compris le développement d'un code d'éthique, est nécessaire pour encadrer le développement et la production des systèmes d'IA. À moyen et long terme, cela complètera de façon réfléchie et informée, le cadre légal existant à l'échelle nationale et européenne. Cette réflexion éthique des ingénieurs doit être fondée sur les principes de bienfaisance ou non-malfaisance, d'autonomie et de justice ainsi que sur les principes et valeurs consacrés dans la Charte européenne des droits fondamentaux.
Pourquoi pas une déontologie des ingénieurs en IA ?

Objectif 1 : "la formation" afin d'augmenter la compréhension de l'éthique par les ingénieurs
Moyen : intégrer l'éthique dans le maquettage des formations
Coûts : ECTS
Évaluations : applicabilité de la formation à l'utilisation des ressources de l'éthique

Objectif 2 : "les ressources de l'éthique" afin d'accompagner la conception de l'IA, et la réévaluation nécessaire

Moyen : rédaction de chartes éthiques, méthodologie d'évaluation, mises en œuvre d'outils pour valider le suivi des chartes dans différents secteurs, par exemple un robot d'aide à la personne, aide au diagnostic, benchmarks sur les capacités des IA
Coûts : ETP + plateformes/réseaux
Évaluations : degré d'utilisation, appropriation (degré), mise à jour

2. Éducation : avec l'arrivée de l'IA, une révolution sociétale est en marche qui nécessite de la part des dirigeants de définir un nouveau projet de société clair et apparent : faire évoluer les formations initiales et l'enseignement dans les écoles et lycées ainsi que les formations aux métiers qui vont être bouleversés.

3. Surveillance et formations : il faut surveiller de près le marché du travail, spécialement la création et la perte d'emplois et de savoir-faire dans différents champs professionnels. Il est fondamental d'identifier le plus tôt possible divers tendances et chiffrer le taux de remplacement par l'IA dans différents secteurs de l'automatisation. **Les comités d'entreprise et les organismes de représentation des travailleurs (syndicats) devront être associés à ces études et formés à ces nouveaux enjeux.**

4. Recherches et études de terrain : il est nécessaire de **tester les dispositifs de l'IA dans des situations réelles ("real-life scenarios") sur le long terme** (ou au moins à plusieurs moments d'utilisation dans le temps) afin d'identifier et d'évaluer les risques qu'ils peuvent engendrer au-delà de la phase d'études à court terme.

5. Implication et information : afin de promouvoir l'acceptabilité sociale des technologies et de faciliter leur appropriation, Il faut **informer les personnes sur les développements technologiques en cours dans la société et les alerter des conséquences potentielles sur leur travail**. Les usagers doivent être adéquatement représentés dans la définition des politiques publiques en matière d'IA et pris en compte dans les processus d'innovation. Les métiers et opportunités de cette révolution ne seront accessibles dans un premier temps qu'aux personnes averties et sensibilisées à ces outils. L'information doit aussi se faire auprès des élèves et étudiants afin de les attirer en amont vers ces formations.

6. Engagement : **dans certains secteurs d'activité, il est important de préserver le contact humain et la dimension affective et corporelle de l'activité**. À ce titre, il faut prévoir d'aménager des « espace-temps » où les interactions sociales sont fortes et où l'implication subjective dans l'activité demeure importante.

7. Prise de décision : il est nécessaire d'amener le travailleur à **coopérer en synergie avec la machine, de le faire participer au processus décisionnel** et de lui conférer un rôle susceptible d'engager sa responsabilité en sollicitant l'engagement personnel et la mobilisation subjective dans le processus de travail. Il faut que le travailleur puisse reprendre le contrôle de la machine (reprise en main) si nécessaire. À cette fin, il est utile de créer de nouveaux métiers tels que « compreneur » ou « interprète » des processus d'apprentissage non supervisé afin de **créer de nouveaux outils que le travailleur pourra utiliser pour connaître l'état de la machine, comprendre son comportement et la surveiller. Il doit aussi pouvoir apporter des suggestions sur des améliorations ou des corrections éventuelles.**

8. Autonomie : préserver l'autonomie et la créativité des individus, notamment en envisageant des possibilités d'improvisation et de spontanéité est une autre priorité. **Dans certains cas, le travailleur doit pouvoir désactiver les dispositifs issus de l'IA** afin d'exercer son activité en toute indépendance et liberté.

9. Intimité et vie privée : il est nécessaire de **garantir la vie privée et une sphère d'intimité dans l'environnement de travail** afin que chacun ait, quand il le souhaite, des moments de répit, propices à la concentration et à la réflexion. Le travailleur et l'utilisateur doivent être informés des données privées collectées ou traitées durant l'utilisation de la machine.

10. Co-travail avec l'IA : dans le co-travail avec une IA, la machine doit respecter les valeurs de l'humain à travers le travail. À cette fin, les travailleurs et les usagers seront formés afin de leur permettre de se familiariser avec les exigences techniques de leurs secteurs et de **développer de nouvelles compétences.** Les standards d'excellence ou de productivité devront probablement être revus pour éviter une compétition perdue d'avance sur des domaines où l'humain n'est pas le plus performant.

11. Normalisation : les autorités doivent lancer une concertation en vue du développement des normes techniques et industrielles pour le déploiement de l'IA. Cela doit comporter des normes spécifiques par secteur industriel ainsi que des standards généraux portant sur les aspects transversaux de l'IA.

12. Expérimentation : La création de laboratoires ou plateformes d'expérimentation des nouveaux métiers, sur l'exemple qui est cité (DIGILABSOCIAL).

Annexe 2

Contribution du sous-groupe 3.2.B

INTELLIGENCE ARTIFICIELLE

ENJEUX JURIDIQUES

13/03/2017

Contribution remise au groupe de travail 3.2 en charge de l'analyse des impacts économiques et sociaux de l'intelligence artificielle

SOMMAIRE

I. Propos préliminaires

En l'absence de cadre juridique spécifique à l'intelligence artificielle (IA), l'objectif de notre groupe de travail fut d'analyser les enjeux principaux en la matière, et d'identifier d'éventuelles solutions, relevant tant du droit « dur » que du droit « souple ». Cet objectif est capital puisqu'il s'agit d'assurer un meilleur contrôle du fonctionnement de l'IA sans pour autant freiner l'innovation.

Plus précisément, notre groupe a décidé de travailler autour de quatre sujets majeurs, liés au développement de l'IA, et porteurs d'enjeux juridiques : **les algorithmes, les données, la responsabilité et la création assistée**. Pour mener à bien nos travaux, nous nous sommes réunis au sein de plusieurs réunions afin d'échanger sur ces sujets, et avons par ailleurs mené des auditions afin de cerner au mieux les attentes des parties prenantes.

Dans le cadre de l'établissement d'une stratégie optimale en matière d'IA, nous pensons qu'il est important de se livrer à un travail préalable de définitions.

En effet, il ressort de nos travaux et des auditions menées que le terme « intelligence artificielle » est souvent utilisé de manière trop englobante, et qu'il convient de rappeler, pour une approche juridique et règlementaire rigoureuse, que cette notion couvre, au sens strict, **les technologies capables de traiter des sources hybrides et notamment des données non structurées**. Elle constitue à ce titre un troisième niveau d'innovation, qui se situe après l'analytics et l'advanced analytics (cette dernière notion couvrant notamment le *machine learning*).

Si les solutions d'intelligence artificielle ne sont pas nouvelles, elles ont connu un réel essor ces dernières années, et ce, en raison de trois facteurs principaux : le volume de données (Big data), des algorithmes de plus en plus sophistiqués (grâce aux progrès informatiques et mathématiques), et l'expansion du cloud.

Les algorithmes et les données sont ainsi au cœur de l'IA : ils en permettent le fonctionnement, et leur relation va mener à son perfectionnement. De manière très simplifiée, on peut considérer qu'un algorithme initial est nourri par des données afin d'aboutir à un algorithme final enrichi. Cette représentation permet de mieux appréhender les questions de propriété et d'accessibilité aux algorithmes et données détenus par une entreprise.

II. L'application limitée des régimes juridiques actuels pour protéger les algorithmes utilisés en matière d'IA

Selon les modèles d'exploitation pratiqués, deux cas de figure se présentent. Dans un premier cas, la technologie d'IA repose, selon le choix de leur exploitant, sur des algorithmes dans le domaine public ou soumis, à titre alternatif, à un régime dit « open source ». C'est ici, la valeur ajoutée résultant de leur combinaison ou de leur éducation/apprentissage que l'exploitant va chercher à protéger. Dans un second cas, c'est la protection des algorithmes en tant que tels qui peut être recherchée. Dans les deux cas, le recours aux droits de propriété intellectuelle (droit d'auteur, logiciel ou brevet) n'est pas totalement satisfaisant et les exploitants semblent privilégier à l'heure actuelle la protection par le secret des affaires. Il ressort des travaux du sous-groupe que la flexibilité (les garanties de confidentialité et conditions d'accès peuvent être définies par

contrat) et la sécurité juridique apportées par cette solution sont accueillies favorablement par les exploitants dans un environnement technologique en évolution constante.

Pour les entreprises qui placent l'IA au centre de leurs activités, se pose la question de la définition d'une stratégie de protection et de valorisation de l'innovation. Il ne s'agit pas ici de la protection du robot ou de la machine en général, mais de la solution innovante placée en son centre, qui lui confère son « intelligence » (logiciel et/ou algorithme). A cet égard, se pose la question de la protection de l'algorithme intégré à la machine. Sur ce sujet, des moyens de protection existent et sont utilisés, même si chacun présentent des limites, et qu'ils ne constituent pas, aujourd'hui, des outils de protection totalement satisfaisants.

L'octroi d'un droit privatif au titre de la Propriété intellectuelle apparait ainsi comme une protection nécessairement indirecte.

Conformément à la théorie de l'unité de l'art, les œuvres de l'esprit sont protégeables par le **droit d'auteur**, quels qu'en soient le genre, la forme d'expression, le mérite ou la destination. En conséquence, le droit d'auteur protège les logiciels ou programmes d'ordinateur, sous réserve du respect des conditions classiques applicables en la matière. En revanche, il est établi que les idées et principes à la base d'une œuvre de l'esprit ne sont pas protégeables par le droit d'auteur. Cette exclusion couvre notamment les méthodes ainsi que les principes mathématiques.

Seule sera donc protégée la mise en forme d'un algorithme, par des documents ou un logiciel (algorithme intégré au code source), à condition toutefois de comporter un certain degré d'originalité. En matière de logiciel, cette originalité s'apprécie selon « la marque de l'apport intellectuel » de son auteur. Contrairement à l'appréciation classique de l'originalité, construite autour de la notion d'empreinte de la personnalité de l'auteur, l'originalité du logiciel est davantage fondée sur un « effort intellectuel personnalisé ». Dès lors, le logiciel intégrant un algorithme devra aller au-delà de la simple exécution. Par ailleurs, la dissociation entre protection du logiciel et de ses fonctionnalités a pour conséquence la possible réutilisation de l'algorithme par un tiers qui serait parvenu à l'extraire légalement.

De son côté, le **droit des brevets** vise précisément à protéger l'innovation, sous réserve de remplir les critères de nouveauté, d'inventivité et d'application industrielle.

Néanmoins, le Code de la propriété intellectuelle exclut expressément les programmes d'ordinateurs et les méthodes mathématiques du champ de la brevetabilité, même si cette interdiction n'a pas vocation à s'appliquer à une invention dont la seule mise en œuvre requiert un programme d'ordinateur. Autrement dit, comme pour le droit d'auteur, la protection indirecte reste envisageable. Une attention particulière doit donc être portée à la rédaction du brevet, celle-ci devant revendiquer le procédé et non le programme en tant que tel. Ainsi, l'Office Européen des Brevets (« OEB ») affirme que « un brevet peut être délivré pour une invention mise en œuvre par un ordinateur qui résout un problème technique de manière non évidente »[1]. Les décisions IBM I et II[2] sont communément citées en la matière afin d'illustrer cette frontière tracée par le droit des brevets, en affirmant qu'un programme d'ordinateur n'est pas exclu de la brevetabilité « si sa mise en œuvre sur un ordinateur produit un effet technique supplémentaire, allant au-delà des interactions physiques "normales" entre programme et ordinateur ». Ainsi, un algorithme ou un logiciel d'IA intégré dans une machine, et dont la mise en œuvre produit un effet technique supplémentaire semblent pouvoir être brevetés. Un avantage est que, contrairement au droit d'auteur, l'algorithme inclus dans la demande de brevet sera protégé au même titre que l'invention elle-même[3]. Toutefois, ce moyen de protection implique de divulguer l'algorithme au public (donc à ses concurrents), et de déposer toute modification de l'algorithme.

Les entreprises auront souvent recours au secret comme palliatif, ou complément, pour une protection efficace de leurs innovations.

La **contractualisation** reste un moyen classique et privilégié par l'entreprise pour protéger ses secrets industriels, que cela passe par un accord ou une clause de confidentialité. Toute divulgation sera ainsi sanctionnée par le droit de la responsabilité contractuelle, et ce même en l'absence de faute. L'intérêt de cette solution réside dans le principe de liberté contractuelle et la créativité offerte aux parties (définition du secret, interdiction d'exploitation, maintien de l'obligation après la cessation du contrat…).

En l'absence de contrat, la voie de la **responsabilité délictuelle** reste envisageable. Ainsi, l'**action en concurrence déloyale** permet de sanctionner des faits de **parasitisme** économique. Toutefois, il faudra être en mesure de démontrer une faute dans l'appropriation et dans la réutilisation du secret[4].

Concernant la voie pénale, il n'existe pas, aujourd'hui, d'infraction permettant de sanctionner directement la violation du secret des affaires. Toutefois, l'algorithme pourra utilement être protégé par la future **Directive 2016/943 du 8 juin 2016 sur le secret des affaires** (qui doit être transposée au plus tard le 9 juin 2018). En effet, cette directive prévoit une définition large du secret des affaires, les informations devant seulement présenter les trois caractéristiques cumulatives suivantes :

- être secrètes « en ce sens que, dans leur globalité ou dans la configuration et l'assemblage exacts de leurs éléments, elles ne sont pas généralement connues des personnes appartenant aux milieux qui s'occupent normalement du genre d'informations en question, ou ne leur sont pas aisément accessibles » ;
- avoir une valeur commerciale du fait qu'elles sont secrètes ;
- avoir fait l'objet, de la part de la personne qui en a le contrôle de façon licite (le « détenteur de secrets d'affaires »), de dispositions raisonnables, compte tenu des circonstances, destinées à les garder secrètes.

L'apport de la directive réside, en outre, dans les mesures de protection et de réparation : mesures conservatoires et provisoires (comme en matière de contrefaçon), procédure spéciale pour préserver la confidentialité des débats lors de la procédure, méthodes particulières de calcul des dommages-intérêts.

Ce nouveau texte doit permettre à l'Union européenne d'assurer une protection efficace du patrimoine immatériel des entreprises. La Directive ne prévoit pas une infraction pénale spécifique, la prérogative étant laissée aux Etats membres. La France pourrait utilement saisir cette occasion pour réprimer l'atteinte au capital immatériel de l'entreprise, dont la valeur n'est plus à démontrer aujourd'hui[5].

III. Transparence et loyauté : un préalable nécessaire au bon développement de l'IA

La confiance dans l'IA est directement liée à la transparence et à la loyauté des traitements algorithmiques. Afin que le besoin légitime de transparence ne devienne pas un frein pour l'innovation et sa valorisation, un équilibre doit être recherché. Aussi, afin de satisfaire cette exigence tant vis-à-vis du consommateur que dans une perspective concurrentielle, des initiatives de co-régulation tendent à l'adoption généralisée de bonnes pratiques. A l'issue de ses travaux, le sous-groupe préconise que la France encourage le développement d'une initiative auprès des acteurs de l'IA pour la mise en place de grands principes de référence en la matière (guides et

bonnes pratiques, norme relative aux critères de transparence…), permettant d'assurer une transparence objective.

A rebours de l'idée de protection et de secret, les questions liées au mouvement de transparence relatif aux traitements opérés et à la preuve de ces traitements constituent également un enjeu juridique pour le développement de l'IA. Le sujet de la transparence est au cœur des débats actuels, en réponse, notamment, aux craintes exprimées envers l'objectivité des algorithmes susceptibles de biais, de discriminations, ou encore de risque pour la diversité de l'information sélectionnée. **Si ces craintes sont, dans une certaine mesure, légitimes, un équilibre doit être préservé entre exigences de transparence des algorithmes et innovation.**

Deux obstacles s'opposent à ces exigences. D'une part, la nature fluctuante et évolutive des algorithmes ; d'autre part, leur valeur d'un point de vue concurrentiel. Comme l'a fort justement souligné le Conseil supérieur de l'audiovisuel (CSA), *« Toute tentative de réguler les algorithmes serait probablement vouée à l'échec en raison de l'évolutivité de la technologie et du caractère confidentiel et concurrentiel des développements. »*[6] En outre, il est important de noter que la transparence et l'objectivité des résultats seront fonction du volume et de la nature des données traitées par l'algorithme. Enfin, il serait contre-productif de « perdre » l'individu dans une masse d'information.

A ce stade, il convient donc surtout de définir ce qu'on entend par le terme de « transparence ». A ce sujet, il existe un relatif consensus sur l'idée de rendre transparent pour l'utilisateur les informations relatives à l'origine des données utilisées par la machine et la finalité poursuivie. Certains principes communs, souples et équilibrés pourraient utilement être adoptés par les concepteurs d'IA. A titre d'exemple, Nicholas Diakopoulos et Sorelle Friedler, de l'Institut de recherche Data & Society, ont proposé cinq principes pour rendre les algorithmes fiables : responsabilité, explicabilité, exactitude, l'auditabilité et justiciabilité[7]. Tout comme en matière de données personnelles, la confiance des utilisateurs dans l'IA (aujourd'hui faible en France[8]) représente un avantage compétitif pour les entreprises. Certains groupes l'ont d'ailleurs bien compris, puisqu'ils commencent à travailler sur ces sujets. A titre d'exemple, Google, Facebook, IBM, Microsoft et Amazon viennent de créer un « Partnership on Artificial Intelligence to Benefit People and Society », en vue de définir ensemble des bonnes pratiques[9]. De son côté, le fondateur de Tesla et SpaceX, Elon Musk, a créé OpenAI[10].

Bien que n'étant pas spécifiques à l'IA, il faut noter que deux outils législatifs trouvent déjà à s'appliquer, et constituent plus largement une potentielle source d'inspiration pour le régulateur. En effet, la **loi pour une République numérique** impose une exigence de transparence aux administrations quant aux traitements algorithmiques servant à prendre des décisions individuelles (art. 4 et 6), et a renforcé l'actuel article L. 111-7 du code de la consommation pour élargir l'exigence de loyauté des plateformes et l'information des consommateurs, au-delà des seules plateformes de mise en relation (art. 49 & suiv.). En application de ce texte, tout opérateur de plateforme en ligne est tenu de délivrer au consommateur une information loyale, claire et transparente sur les conditions générales d'utilisation du service d'intermédiation qu'il propose, et *« sur les modalités de référencement, de classement et de déréférencement des contenus, des biens ou des services auxquels ce service permet d'accéder »*. Par ailleurs, les opérateurs de plateformes en ligne dont l'activité dépasse un seuil de nombre de connexions devront élaborer et diffusent aux consommateurs des bonnes pratiques visant à renforcer les obligations de clarté, de transparence et de loyauté.

De son côté, le **règlement général de l'UE sur la protection des données personnelles (« RGPD »)** prévoit que le responsable de traitement doit informer la personne concernée de *« l'existence d'une prise*

de décision automatisée, y compris un profilage[11] *(…) et, au moins en pareils cas, des informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée »* (art. 13-2-f). A partir de cette information, est offert à la personne concernée un droit d'opposition[12]. Deux exceptions sont néanmoins prévues à cette faculté : lorsque cela *« est nécessaire à la conclusion ou à l'exécution d'un contrat entre la personne concernée et un responsable du traitement »* ; ou si le consentement explicite de la personne a été recueilli[13].

On voit que, quel que soit le vocable utilisé, l'idée sous-jacente est toujours de permettre à l'individu de comprendre le fonctionnement de l'IA. On retrouve cette idée dans la récente **résolution du Parlement européen sur la robotique**[14], qui *« insiste sur le principe de transparence, à savoir qu'il devrait toujours être possible de fournir la justification rationnelle de toute décision prise avec l'aide de l'intelligence artificielle »*[15]. A cet effet, le Parlement européen propose que chaque robot soit doté d'une « boîte noire » contenant les données sur chaque opération réalisée par la machine, y compris les logiques ayant contribué à la prise de décisions.

Il est important de préciser que la question de l'éthique ne doit pas se limiter aux algorithmes mais doit aussi couvrir les conditions d'accès aux données, tant personnelles que non personnelles ou anonymisées. En effet, en matière d'IA, algorithmes et données forment un tout indissociable, à l'image d'un moteur et de son carburant. Comme l'a expliqué l'entrepreneur Rand Hindi, *« La donnée est à l'Intelligence Artificielle ce que l'uranium est au nucléaire, avec la même problématique des usages qui en sont faits (civils ou militaires). »*[16] En ce qui concerne les données personnelles, il ne fait aucun doute que la conception et la mise en œuvre de l'IA se doivent d'être **respectueuses de la vie privée et des données personnelles**, sachant que cette dernière notion est très large, et l'anonymisation rarement irréversible. L'IA ne bénéficiant pas encore d'une entière autonomie, il revient aux concepteurs de garantir un niveau élevé de protection.

A cet égard, le RGPD fournit aux acteurs un cadre modernisé pour développer leurs produits et services. Les nouveaux principes qu'il consacre, notamment ceux de *privacy* et de *security by design*, ou encore d'*accountabilty*, trouvent ici toute leur pertinence, même si on ne peut ignorer qu'ils peuvent techniquement constituer un frein ou une barrière à l'entrée de certains acteurs émergents. D'autres règles, plus classiques, doivent également être respectées, comme la minimisation des données et la limitation des finalités. S'il ne semble donc pas nécessaire, à ce stade, de s'orienter vers une régulation spécifique relative à la protection des données en matière d'IA, certains concepts mériteraient davantage d'éclaircissement afin d'éviter toute source d'insécurité juridique, tant pour les entreprises que les individus. A ce titre, il est primordial que soit rapidement adopté un code de conduite ou un mécanisme de certification, pour démontrer le respect des obligations incombant au responsable du traitement ainsi qu'aux sous-traitants. Sur le sujet de la sécurité, la Commission européenne a indiqué qu'elle présenterait, d'ici fin 2017, une proposition de mécanismes de certification et de labélisation communs en matière d'IoT afin d'assurer la cybersécurité de ces dispositifs. Une telle intervention ne doit cependant pas freiner l'innovation. Un niveau élevé de sécurité des systèmes reste, néanmoins, indispensable à une utilisation appropriée de l'IA. Les acteurs doivent donc concevoir des produits sûrs, sécurisés et adaptés à l'utilisation à laquelle ils sont destinés. Toutefois, on peut supposer que les acteurs appliqueront un principe de proportionnalité afin d'assurer un certain équilibre sécurité/coûts/risques. A cet égard, l'harmonisation des normes techniques doit être poursuivie au niveau européen et international, en gardant à l'esprit que leur traduction technique peut s'avérer délicate dans la mise en œuvre (temps de traitement par ex.) ou tout simplement coûteuse, générant ainsi une possible barrière à l'entrée pour certains acteurs.

IV. Propriété des données : préserver les investissements sans bloquer l'accès aux données nécessaires au développement de l'IA

S'agissant des modèles économiques et techniques que le sous-groupe a eu l'occasion d'examiner dans le cadre de ses travaux, il apparait que l'exploitant d'IA est positionné contractuellement vis-à-vis de son client en tant que prestataire de solution technique. Tant les données brutes, objets du traitement algorithmique, que les données résultant de ce traitement, demeurent contractuellement la propriété du client. Dans un tel schéma, la problématique de propriété des données ne semble pas poser de difficulté particulière, pour autant que l'équilibre des intérêts soit de mise. Afin d'assurer cet équilibre, une veille sur les clauses contractuelles utilisées par les opérateurs économiques pourrait être mise en place afin d'identifier et de développer des bonnes pratiques en la matière.

Dans son étude sur le numérique et les droits fondamentaux, le Conseil d'Etat s'opposait à la reconnaissance d'un droit de propriété des individus sur leurs données, écartant par-là l'introduction d'une logique patrimoniale dans la protection des droits personnelles. En effet, la protection des données personnelles, en France et en Europe, repose sur une logique de droits attachés à la personne. Pour cette raison, le Conseil soutenait davantage la consécration du droit à l'autodétermination, notion qui empreint indirectement la loi pour une République numérique et le RGPD. Ainsi si les individus ne disposent pas d'un titre de propriété sur leurs données personnelles, celles-ci restent, en théorie, entre leurs mains et ils doivent pouvoir en disposer dans le cadre de leur droit à la portabilité[17].

Au-delà des données à caractère personnel, se pose aujourd'hui avec acuité **la question de la propriété des données non-personnelles ou anonymisées**. L'enjeu ici est de trouver un équilibre entre la valorisation économique de ce « nouvel or noir » et l'accessibilité des données (*open data)*, nécessaire à l'innovation.

En France, la loi pour une République numérique a ouvert les données publiques, en ce compris les données d'intérêt général (délégations de service public ou subventions, énergie…. Elle n'est, en revanche, pas intervenue sur la question de l'accessibilité à d'autres types de données.

La Commission européenne s'interroge actuellement sur le cas des entreprises qui détiennent des données « brutes » (i.e. données qui n'ont pas été traitées ou modifiées depuis leur enregistrement) et anonymes, et les enjeux concurrentiel associés. Il est important de noter que la position française au niveau européen est très en soutien de cette l'initiative de la Commission en matière de Free Flow Of Data[18]. La Commission porte, à cet égard, une attention particulière aux données collectées par des capteurs intégrés dans des machines, des outils et/ou des dispositifs (ex. IoT), et qui sont en mesure de les partager[19]. Afin de favoriser l'innovation et d'éviter les exploitations en silos, la Commission soumet à la réflexion plusieurs options pour améliorer l'accès à ces données, dont entre autres : l'élaboration de lignes directrices, l'adoption de clauses contractuelles par défaut, la généralisation d'un contrôle des clauses abusives B-to-B, ou encore le développement de formats ouverts et d'APIs. Le Parlement européen rejoint ces objectifs d'ouverture et d'accessibilité dans sa résolution sur la robotique[20].

Au-delà des lignes directrices et des clauses encouragées par les institutions, il ressort des auditions et contributions qui ont alimenté les présents travaux, que le mode de développement et de fonctionnement de l'IA, organisé en écosystèmes très évolutifs, peut avantageusement être aménagé autour de contrats. Conclus entre les différents acteurs de l'IA (plateformes, réseaux, fournisseurs d'algorithmes ou de

données), les contrats permettent, au plus près de la réalité technique, de définir leurs rôles et leurs responsabilités dans le développement et dans le rendu des services créés autour de l'IA, et d'aménager les conséquences qui en résultent pour les utilisateurs ou les bénéficiaires finaux. Protecteur des intérêts en balance (créateurs, bénéficiaires, utilisateurs), le contrat, par la liberté et la souplesse qu'il offre, permet de s'accommoder, à juste temps, des situations d'agilité nécessaire aux démarches permanentes de recherche et développement, d'innovation et d'amélioration dans lesquelles s'inscrit l'IA enrichie en continue par le *machine learning*. Le contrat se présente donc comme une réponse pour l'appréhension de cette technologie « fugitive » que constitue l'IA[21] : s'il demeure un lieu de concordance des volontés des acteurs, il permet aussi de refléter, au-delà de leur commune intention, les intérêts des parties. A cet égard, il s'agit conformément à notre droit commun des contrats, de protéger la partie affaiblie dans le processus de négociation et de rendu des services de l'IA, afin de la prémunir contre tout abus. Mais il s'agit aussi, à la recherche d'un équilibre contractuel toujours fragile, de préserver la viabilité des modèles techniques et économiques émergents avec l'IA. Afin d'assurer cet équilibre, une veille sur les clauses contractuelles utilisées par les opérateurs économiques pourrait être mise en place afin d'identifier et de développer des bonnes pratiques en la matière.
C'est dans cet esprit que la Commission européenne compte aussi évaluer prochainement la **Directive 96/9/CE** du 11 mars 1996 concernant la protection juridique des bases de données, et interroge sur la question de la propriété des données brutes, avec en vue l'objectif d'accroitre l'octroi de licences d'utilisation. A cet égard, si elle envisage un partage de la propriété (entre le propriétaire, l'utilisateur et le fabricant), elle propose aussi la reconnaissance d'un droit *sui generis* de propriété intellectuelle au fabricant ou au producteur.

En effet, aujourd'hui, seul le producteur d'une base de données bénéficie d'un droit de propriété, lorsque la constitution de la base « atteste d'un investissement financier, matériel ou humain substantiel » (art. L.341-1 du code de la propriété intellectuelle). Ce droit *sui generis* vient en supplément du droit d'auteur qu'il complète. A contrario, les investissements pour la création de données ne sont pas pris en compte, ni la gestion de données et les offres de services sans constitution de bases permanentes. Une grande partie des traitements de données récents et innovants (ex. data mining) ne sont donc pas couverts par le régime de la loi du 1er juillet 1998. Des auteurs proposent ainsi de l'étendre en un « droit des producteurs et des exploitants de données et de bases de données »[22]. Cette extension impliquant au demeurant d'identifier *« quand et pourquoi le lien entre le producteur de bases de données et ses données pourra se rompre, au profit du curateur ou de celui qui l'emploie »*, ce qui pourrait résider dans la condition première d'un effort économique (au sens large) propre. Comme le soulignait Nicolas Courtier, avocat, il s'agit de permettre à des activités de traitement et d'analyse de cohabiter[23].

Enfin, concernant l'exception de *text and data mining*, introduite en France par la Loi pour une République numérique, à la suite notamment du Royaume-Uni et figurant dans le projet de réforme européenne du droit d'auteur, certains acteurs craignent qu'en recentrant l'accès sur la seule recherche, la fouille numérique de textes et de données soit *de facto* exclue de leurs licences en l'absence de disposition expresse. Il semble donc nécessaire de prévoir expressément cette hypothèse contractuellement.

V. La responsabilité du fait des objets intelligents

Il résulte des travaux du sous-groupe qu'il n'y pas à l'heure actuelle de vide juridique, étant donné que les instruments du droit positif de la responsabilité peuvent être mobilisés, afin de trancher les litiges susceptibles d'émerger. Seul le recul de l'expérience pratique, au double point de vue

juridique et économique, permettra, à terme, de faire des choix. Ce recul permettra de proposer, le cas échéant, une adaptation pertinente en termes de droit de la responsabilité, si les limites pressenties des régimes actuels devaient être confirmées. Des règles spéciales doivent cependant être d'ores et déjà envisagées s'agissant de certaines machines intelligentes, en particulier les voitures autonomes.

La réception par le droit du développement de la robotique et de l'IA pose au premier rang la question de la responsabilité du fait des objets intelligents. La résolution adoptée par le Parlement européen le 16 février 2017 « *contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique* » l'accompagne d'une mise en perspective que serait une responsabilité des robots eux-mêmes, ce qui oriente la réflexion vers la reconnaissance d'une **personnalité juridique** de ces machines intelligentes. Cette conception doit cependant être repoussée car elle est aussi inutile que dangereuse. En effet, la personnalité attribuée à certains objets intelligents – lesquels ?, selon quel critère ? – ne règlerait rien des questions qui se posent en termes de sécurité et de responsabilité. La sécurité peut plus sûrement compter sur des normes imposées aux concepteurs de logiciels, aux fabricants, voire aux utilisateurs de robots qu'en programmant des devoirs dont ces derniers seraient personnellement débiteurs. Quant à la responsabilité, elle se conçoit rationnellement comme une responsabilité du fait de la chose robot plutôt que comme une responsabilité du robot lui-même qui impliquerait d'abonder un patrimoine qui lui serait attaché pour l'exécution de la dette de réparation. La personnification des robots troublerait surtout gravement les catégories juridiques en donnant vie à une chimère, mi-personne mi-chose, à la fois sujet de droit et objet de droit.

A s'en tenir à une responsabilité du fait du robot ou de l'IA lorsqu'elle n'est pas intégrée dans une enveloppe physique, la détermination d'un responsable, ou du moins d'un débiteur d'indemnités par le truchement de l'assurance, peut se faire au sein d'un **groupe de responsabilités particulièrement large** : propriétaire, utilisateur, développeur de l'IA (ou sans doute plutôt éditeur), fabricant, etc. En pratique, la victime pourra attraire l'ensemble de ces intervenants, lesquels pourront le cas échéant opposer les uns aux autres des clauses limitatives ou exonératoires de responsabilité stipulées dans les contrats qui les lient. **En l'état actuel des choses, il nous semble prématuré de vouloir profiler un régime plutôt qu'un autre**. Les choix ne pourront être faits, au double point de vue juridique et économique, qu'avec **le recul de l'expérience pratique**. Il faut, en particulier, compter sur les orientations que prendront les systèmes assurantiels, lesquelles détermineront le point d'ancrage de la couverture du risque. La résolution adoptée par le Parlement européen est sur cette ligne de temporalité, considérant que de nouveaux instruments ne pourront être mis en place qu'à un horizon de dix à quinze ans (point 51). Durant la période intermédiaire, il n'y a pas de vide juridique car **les instruments du droit positif peuvent être mobilisés**. Dans cette optique d'un traitement de la réparation par le droit actuel, la nature des risques devrait être déterminante dans le choix du mécanisme de responsabilité : si les risques de structure liés à la conception de l'objet intelligent peuvent justifier la responsabilité spéciale du fabricant du fait des produits défectueux, les risques de comportement tenant à l'utilisation de la chose entraîneraient la responsabilité de droit commun du fait des choses du propriétaire ou de l'utilisateur à l'égard de tiers. Tout au plus, la pratique révèlera au besoin la nécessité d'adapter certains éléments de régime comme l'exonération pour risque de développement attachée à la responsabilité du fait des produits défectueux.

Sans attendre, des règles spéciales pourraient d'ores et déjà être adoptées en ce qui concerne certaines machines intelligentes. C'est le cas, en particulier, des **véhicules autonomes**, car il fait peu de doute que leur mise sur le marché suppose la confection d'une règlementation qui sécurise l'utilisation de ces nouveaux objets, et l'indemnisation des dommages qu'ils sont susceptibles de causer. Ne serait-ce que d'un point de vue d'acceptabilité sociale, la norme doit anticiper la mise en pratique. Un travail

préparatoire a d'ailleurs déjà commencé à propos de leur utilisation. Par ordonnance du 3 août 2016 prise en application de la loi **n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte,** le gouvernement a autorisé la circulation sur la voie publique de véhicules autonomes pour l'expérimentation de ces voitures sans conducteur ou disposant de systèmes d'aides à la conduite. **Le texte constitue un préalable à l'élaboration d'un cadre règlementaire par décret en Conseil d'Etat.**

Dans une vision plus générale de la responsabilité, qui ne doit pas se limiter pas à une question de réparation de dommages, **l'attention doit être attirée sur deux points**. Le premier est l'asymétrie des responsabilités qui se profile entre celle du fabricant de robots incarnés intégrant une IA et celle du développeur/éditeur d'IA. La responsabilité du fabricant du robot physique (la partie *hard*) intégrant une intelligence artificielle (la partie *soft* dont le fabricant n'est pas le plus souvent propriétaire) pourra être assez facilement engagée en cas de dommage, le *hard* et le *soft* formant un tout, un produit complexe. En revanche, la responsabilité de l'éditeur sera plus difficile à établir. Il ressort des auditions que les d'IA mettent contractuellement à disposition du client un simple outil, le système d'apprentissage dépendant de l'utilisateur-client qui ainsi contribue au processus décisionnel. L'autre point est la possibilité de risques systémiques. Le développement de l'intelligence artificielle entraînera des situations de « captivité » dans lesquelles certaines activités dépendront pour leur fonctionnement d'applications de l'IA. De quelles ressources dispose-t-on si l'IA ne fonctionne plus ou ne fonctionne plus correctement ? L'impact économique peut dépasser l'approche sectorielle et entamer la souveraineté (numérique) nationale.

Ces deux considérations doivent conduire à réfléchir à la place et au rôle du développeur dans la chaîne. Il ne s'agit pas d'ajouter un régime de responsabilité et de prendre le risque de brider le développement économique du secteur, mais des obligations pourraient être imposées sous forme de principes au développeur d'IA afin de rééquilibrer l'asymétrie des responsabilités et de prévenir les risques systémiques. On peut penser, en premier lieu, au principe de transparence précité, notamment par l'explication de protocoles de test et des résultats obtenus sur chaque itération d'IA ou encore l'obligation d'indiquer l'objectif poursuivi par l'algorithme prédictif – ce qui peut être rapproché du principe, précédemment évoqué, de finalité du traitement en matière de données personnelles ou de la transparence exigée dans les algorithmes décisionnels publics par la loi pour une République numérique (v. art. L. 311-3-1 du Code des relations entre le public et l'administration). Ce principe de transparence pourrait s'accompagner d'un principe de continuité des services décidée par une autorité administrative indépendante du type de l'agence européenne chargée de la robotique et de l'intelligence artificielle dont la création est préconisée dans la résolution adoptée par le Parlement européen. D'autres principes peuvent être envisagés : principe de neutralité de la base de référence, des tests et des traitements algorithmiques pour éviter d'aboutir à des résultats discriminatoires notamment ; principe de séparation des IA pour limiter le risque de propagation de biais viciant une intelligence artificielle ; droit de s'opposer à un traitement algorithmique ou de contester une décision issue d'un traitement algorithmique…

VI. Propriété intellectuelle et IA

L'IA peut intervenir dans le processus de création littéraire et artistique à deux titres principaux. Dans un premier cas de figure, l'IA peut être présentée comme un outil d'aide à la création, dont l'objet est de fournir une assistance à une personne physique. Les solutions juridiques traditionnelles sont applicables à cette situation, qui aboutit à l'attribution de la titularité des droits à la personne physique. Un second cas de figure correspond à la situation où l'IA, de manière autonome, développerait un résultat susceptible d'être qualifié d'œuvre, sous réserve du critère de l'originalité. Toutefois, en l'absence d'intervention d'un auteur personne physique, le critère

d'originalité ne semble pas pouvoir être rempli, sauf à procéder à une refonte totale du régime du droit d'auteur. Selon le sous-groupe, la recherche des solutions juridiques les plus adaptées à ce dernier cas de figure devrait faire l'objet d'une réflexion approfondie dans le cadre de la stratégie à venir.

Les créations qui seraient générées par une intelligence artificielle, sans intervention directe d'une personne physique dans le processus créatif du résultat considéré risquent d'échapper à l'appropriation par le droit d'auteur, qui nécessite de qualifier l'empreinte de la personnalité ou encore l'effort technique personnalisé. Le droit d'auteur est intimement lié à la personne humaine dont la subjectivité et la capacité créative détermine l'accession à la protection.

Les œuvres assistées par ordinateur sont protégeables par le droit d'auteur, l'ordinateur appréhendé comme un outil n'étant pas exclusif de créativité humaine. Ainsi, lorsque le processus créatif est laissé à la personne physique qui apporte l'empreinte de sa personnalité, cette personne sera auteur, au sens du Code de la propriété intellectuelle, de l'œuvre considérée sous réserve de son originalité. A ce titre, la Cour d'appel de Paris a déjà jugé que l'œuvre assistée par ordinateur « *peut être protégée par le droit d'auteur à conditions qu'apparaisse l'originalité voulue par le concepteur* » (Cour d'appel de Paris, 3 mai 2006, RG 05/03736). Tel sera le cas de la création générée par une intelligence artificielle au moyen des données sélectionnées par l'utilisateur en considération de ses choix et de sa personnalité. Sur un même modèle, la Nouvelle-Zélande, dans son Copyright Act du 15 décembre 1994, envisage la protection des œuvres créées par une intelligence artificielle dès lors que le processus créatif relève des choix d'une personne physique (article 5 : "*the person who creates a work shall be taken to be, (a) in the case of a literary, dramatic, musical, or artistic work that is computer-generated, the person by whom the arrangements necessary for the creation of the work are undertaken*").

Tel n'est pas le cas de l'œuvre réalisée de façon autonome par une intelligence artificielle bénéficiant d'une capacité d'analyse de l'environnement ainsi que de facultés d'apprentissage et de subjectivité. La raison de cette absence de protection réside dans la définition de l'originalité qui est intimement liée à la personnalité de l'auteur et à l'intervention humaine sans laquelle aucune création ne peut prétendre à la protection par le droit d'auteur. L'importance de ce lien entre l'humain et la création n'est pas spécifique à la France : la Cour suprême australienne et la Cour de cassation hollandaise ont rappelé leur attachement à l'intervention humaine[24] et le Copyright Office américain en a fait une condition de l'enregistrement depuis l'affaire qui avait opposé un photographe et Wikimedia, le premier réclamant au second des droits sur un autoportrait réalisé avec son appareil par un singe[25].

Sur ce sujet, si certains auteurs estiment qu'il ne peut y avoir de droit d'auteur en l'absence de création réalisée par une personne physique d'autres au contraire prônent une objectivisation du droit d'auteur (tant quant aux conditions d'accès à la protection qu'au régime d'exercice de ce droit). Le modèle du Royaume-Uni est alors cité en exemple en ce qu'il prévoit dans le Copyright, Design and Patent Act de 1988 la dévolution des droits sur les créations générés par ordinateur (article 178 (b) : "*is computer-generated a work which is generated by computer in circumstances where there is no human author, in connection with a work*") à la personne ayant procédé aux paramétrages nécessaires à la création (article 9.3 "*In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person who undertakes the necessary arrangements for the creation of the work*"). Par ailleurs, le projet de rapport contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique du 31 mai 2016 (2015/2103 (INL)), dit Rapport Delvaux, « *relève qu'il n'existe aucune disposition juridique qui s'applique spécifiquement à la robotique, mais que les régimes et doctrines juridiques existants peuvent s'appliquer en l'état à ce domaine, certains aspects nécessitant néanmoins un examen approfondi* ». Le rapport « *demande de définir des critères de "création*

intellectuelle propre" applicables aux œuvres protégeables par droit d'auteur créées par des ordinateurs ou des robots ». Cette recommandation n'a pas été retenue dans le rapport final adopté par le Parlement européen le 16 février 2017.

Il résulte des échanges du Sous-groupe que l'exclusion des créations générées par une intelligence artificielle autonome pourrait avoir un impact sur la valorisation des actifs immatériels au sein de l'entreprise notamment et nécessiterait de s'interroger sur une stratégie de valorisation adéquate. Toutefois, une telle stratégie de valorisation ne passerait pas obligatoirement par une refonte du droit d'auteur en niant le lien entre l'œuvre et l'humain, mais pourrait le cas échéant donner lieu à un régime de protection distinct.

***

A titre de constat général, le sous-groupe observe que les initiatives de régulation de l'IA, pour autant qu'elles se révèleraient nécessaires, ne devraient intervenir qu'une fois une certaine expérience acquise et ce, afin d'éviter de brider l'innovation. Dans cette perspective, l'innovation réglementaire aurait toute sa place et la France pourrait privilégier, au côté de la co-régulation, une initiative de règlementation expérimentale telle que prévue à l'article 37-1 de la Constitution[26]. Ce droit à l'expérimentation apparaît un outil intéressant pour appréhender un domaine d'innovation forte telle que l'IA. Il pourrait permettre, par exemple, de tester la mise en place de réglementations dans un secteur donné.

Annexe n°1 – Synthèse des recommandations

Recommandation n°1 : Selon les modèles d'exploitation pratiqués, deux cas de figure se présentent concernant la protection juridique de l'IA. Dans un premier cas, la technologie d'IA repose, selon le choix de leur exploitant, sur des algorithmes dans le domaine public ou, à titre alternatif, soumis à un régime dit « open source ». C'est ici, la valeur ajoutée résultant de leur combinaison ou de leur éducation/apprentissage que l'exploitant va chercher à protéger. Dans un second cas, c'est la protection des algorithmes en tant que tels qui peut être recherchée. Dans les deux cas, le recours aux droits de propriété intellectuelle (droit d'auteur, logiciel ou brevet) n'est pas totalement satisfaisant et les exploitants semblent privilégier à l'heure actuelle **la protection par le secret des affaires**. Il ressort des travaux du sous-groupe que la flexibilité (les garanties de confidentialité et conditions d'accès peuvent être définies par contrat) et la sécurité juridique apportées par cette solution sont accueillies favorablement par les exploitants dans un environnement technologique en évolution constante *(Voir p. 2).*

Recommandation n°2 : La confiance dans l'IA est directement liée à la transparence et à la loyauté des traitements algorithmiques. Afin que le besoin légitime de transparence ne devienne pas un frein pour l'innovation et sa valorisation, un équilibre doit être recherché. Aussi, afin de satisfaire cette exigence tant vis-à-vis du consommateur que dans une perspective concurrentielle, des initiatives de co-régulation tendent à l'adoption généralisée de bonnes pratiques. **A l'issue de ses travaux, le sous-groupe préconise que la France encourage le développement d'une initiative auprès des acteurs de l'IA pour la mise en place de grands principes de référence en la matière (guides et bonnes pratiques, norme relative aux critères de transparence…), permettant d'assurer une transparence objective.** *(Voir p. 4).*

Recommandation n°3 : S'agissant des modèles économiques et techniques que le sous-groupe a eu l'occasion d'examiner dans le cadre de ses travaux, il apparait que l'exploitant d'IA est positionné contractuellement vis-à-vis de son client en tant que prestataire de solution technique. Tant les données brutes, objets du traitement algorithmique, que les données résultant de ce traitement, demeurent contractuellement la propriété du client. Dans un tel schéma, **la problématique de propriété des données ne semble pas poser de difficulté particulière, pour autant que l'équilibre des intérêts soit de mise. Afin d'assurer cet équilibre, une veille sur les clauses contractuelles utilisées par les opérateurs économiques pourrait être mise en place afin d'identifier et de développer des bonnes pratiques en la matière.** *(Voir p. 7).*

Recommandation n°4 : Concernant la responsabilité juridique des objets intelligents, il résulte des travaux du sous-groupe qu'il n'y pas à l'heure actuelle de vide juridique, étant donné que les instruments du droit positif de la responsabilité peuvent être mobilisés, afin de trancher les litiges susceptibles d'émerger. **Seul le recul de l'expérience pratique, au double point de vue juridique et économique, permettra, à terme, de faire des choix**. Ce recul permettra de proposer, le cas échéant, une adaptation pertinente en termes de droit de la responsabilité, si les limites pressenties des régimes actuels devaient être confirmées. **Des règles spéciales doivent cependant être d'ores et déjà envisagées s'agissant de certaines machines intelligentes, en particulier les voitures autonomes** *(Voir p. 9).*

Recommandation n°5 : L'IA peut intervenir dans le processus de création littéraire et artistique à deux titres principaux. Dans un premier cas de figure, l'IA peut être présentée comme un outil d'aide à la création, dont l'objet est de fournir une assistance à une personne physique. Les solutions juridiques traditionnelles sont applicables à cette situation, qui aboutit à l'attribution de la titularité des droits à la personne physique. Un second cas de figure correspond à la situation où l'IA, de manière autonome, développerait un résultat susceptible d'être qualifié d'œuvre, sous réserve du critère de l'originalité. Toutefois, en l'absence d'intervention d'un auteur personne physique, le critère d'originalité ne semble pas pouvoir être rempli, sauf à procéder à une refonte totale du régime du droit d'auteur. **Selon le sous-groupe, la recherche des solutions juridiques les plus adaptées à ce dernier cas de figure devrait faire l'objet d'une réflexion approfondie dans le cadre de la stratégie à venir.** *(Voir p. 10).*

Recommandation n°6 : A titre de constat général, le sous-groupe observe que les initiatives de régulation de l'IA, pour autant qu'elles se révèleraient nécessaires, ne devraient intervenir qu'une fois une certaine expérience acquise et ce, afin d'éviter de brider l'innovation. Dans cette perspective, l'innovation réglementaire aurait toute sa place et la France pourrait privilégier, au côté de la co-régulation, une initiative de règlementation expérimentale telle que prévue à l'article 37-1 de la Constitution[27]. Ce droit à l'expérimentation apparaît un outil intéressant pour appréhender un domaine d'innovation forte telle que l'IA. Il pourrait permettre, par exemple, de tester la mise en place de réglementations dans un secteur donné.

Annexe n°2 – Membres du sous-groupe et auteurs

Georgie COURTOIS - Avocat Associé, De Gaulle Fleurance & Associés / Responsable du groupe

Alexandra BENSAMOUN - Professeur, Université Rennes / CERDI Paris XI / Co-Responsable du groupe

Danièle BOURCIER - Directrice de recherche CNRS, CERSA (CNRS)

Sébastien DUPLAN - Délégué aux relations institutionnelles, Syntec Numérique

Stéphane LARRIERE - Directeur juridique, ATOS

Grégoire LOISEAU - Professeur, Université Paris I Sorbonne

Jean-Sébastien MARIEZ - Senior counsel, De Gaulle Fleurance & Associés

Marc MOSSE – Vice-président de l'AFJE en charge de la transition numérique, AFJE

Camille PARA - Responsable affaires publiques, TECH IN France (ex AFDEL)

Marie SOULEZ - Avocat, Alain Bensoussan Avocats Lexing

Danièle BOURCIER - Directrice de Recherches, C.N.R.S. CERSA

Olivier GUILHEM – Directeur juridique, Softbank robotics

Nina GOSSE – Elève-avocate stagiaire, De Gaulle Fleurance & Associés

Annexe n°3 – Auditions et contributions

Dans le cadre de ses travaux, notre sous-groupe a organisé la tenue des auditions suivantes :

v **<u>Mercredi 1er mars 2017</u> : auditions des représentants de MICROSOFT et d'YSEOP**

Bernard OURGHANLIAN - Directeur Technique et Sécurité, MICROSOFT

Corinne CAILLAUD - Directrice Affaires Extérieures, Affaires Publiques et Juridiques, MICROSOFT

Jean-Renaud ROY - Directeur des Affaires Publiques, MICROSOFT

Mathieu COULAUD - Head of Legal, MICROSOFT

John RAUSCHER - Directeur general, YSEOP

- **<u>Mercredi 8 mars 2017</u> : audition des représentants d'IBM et de SAP**

Vincent MARTINO - Direction juridique, IBM

Diane DUFOIX-GARNIER - Directrice affaires publiques, IBM

Amal TALEB - Directrice adjointe affaires publiques, SAP

- **<u>Lundi 13 mars 2017</u> : audition des représentants des représentants de GOOGLE**

Nathan KLEIN - Juriste

Olivier ESPER - Senior Manager Public Policy

Hugo RUGGIERI - Stagiaire au département contentieux

- **Contributions écrites**

Benoît GOURDON – CEO, TELLMEPLUS (membre de Tech In France)

Pascal GAUVRIT - Chief Technology Officer, EPTICA (membre de Tech In France)

Riccardo MASUCCI - Senior Privacy and Security Policy Manager, INTEL

Mario ROMAO, Senior Manager Policy, INTEL

Lionel THOUMYRE - Director, Public Policy in IP, Content & Media EMEA), INTEL

---

[1] Des brevets pour les logiciels ? Droit et pratique en Europe, EPO, février 1999.

[2] OEB Ch. des recours techniques IBM I T 1173/97 du 1-7-1998 ; IBM II T 935/97 du 4-2-1999.

[3] Quelle protection juridique pour l'algorithme ?, Marc Schuler et Benjamin Znaty, Propriété Intellectuelle & la transformation numérique de l'économie, INPI, 2015.

[4] La Cour d'appel de Versailles a ainsi reconnu comme agissement parasitaire fautif le fait que « la société ayant bénéficié du travail d'analyse, des algorithmes et des codes sources du programme […] même s'il n'est pas contestable qu'elle a procédé à une réécriture complète de son propre logiciel » (Cour d'appel de Versailles, 12ème chambre, section 2, Arrêt du 9 octobre 2003, Microsoft France / Synx Relief et autres).

[5] La proposition de loi de MM. Bruno le Roux et Jean-Jacques Urvoas, relative à la protection du secret des affaires, déposée à l'Assemblée nationale le 16 juillet 2014, proposait notamment une telle infraction.

[6] Pour Dominique Cardon, deux choses sont essentielles et devraient être rendues publiques : savoir ce qui rentre dans la machine (quelles données ?) et quel est l'objectif d'apprentissage inculqué (cf. Compte-rendu de l'événement, organisé par la Cnil, de lancement du cycle de débats publics sur les enjeux éthiques des algorithmes, 23 janvier 2017.

[7] How to Hold Algorithms Accountable, Nicholas Diakopoulos and Sorelle Friedler, Technology Review, 17 novembre 2016. En matière audiovisuelle, le CSA utilise la notion de « loyauté », et propose la publication d'indices mesurant différents aspects qualitatifs des algorithmes (cf. Le rôle des données et des algorithmes dans l'accès aux contenus, CSA, janvier 2017).

[8] Sondage CNIL/IFOP, Notoriété et attentes vis-à-vis des algorithmes, janvier 2017.

[9] https://www.partnershiponai.org/

[10] https://openai.com/blog/

[11] L'article 4 du RGPD définit le profilage comme « toute forme de traitement automatisé de données à caractère personnel consistant à utiliser ces données à caractère personnel pour évaluer certains aspects personnels relatifs à une personne physique, notamment pour analyser ou prédire des éléments concernant le rendement au travail, la situation économique, la santé, les préférences personnelles, les intérêts, la fiabilité, le comportement, la localisation ou les déplacements de cette personne physique ».

[12] L'article 22 dispose ainsi que « La personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire. »

[13] On relèvera également l'article 39 de l'actuelle loi Informatique et Libertés selon laquelle « Toute personne physique justifiant de son identité a le droit d'interroger le responsable d'un traitement de données à caractère personnel en vue d'obtenir (…) 5° Les informations permettant de connaître et de contester la logique qui sous-tend le traitement automatisé en cas de décision prise sur le fondement de celui-ci et produisant des effets juridiques à l'égard de l'intéressé (…) »).

[14] Règles de droit civil sur la robotique, Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique.

[15] Ibid, Considérant n°12.

[16] Compte-rendu de l'événement de lancement du cycle de débats publics sur les enjeux éthiques des algorithmes, CNIL, 23 janvier 2017.

[17] Voir sur ce point les recommendations du G29 sur la portabilité dés donnnées dans le cadre du RGPD : http://ec.europa.eu/information_society/newsroom/image/document/2016-51/wp242_en_40852.pdf

[18]Voir sur ce point : http://www.economie.gouv.fr/files/files/PDF/NAFlibrecirculationdesdonnees_20161110.pdf

[19] Communication « pour créer une économie européenne fondée sur les données », 10 janvier 2017.

[20] Le Parlement européen demande notamment à la Commission de « promouvoir un environnement ouvert qui passe notamment par des normes et des plateformes ouvertes, des modèles d'attribution des licences innovants et de la transparence, afin d'éviter les effets de verrouillage dans les systèmes exclusifs qui limitent l'interopérabilité » (considérant n°9).

[21] V. N. Colin et H. Verdier l'âge de la multitude, Armand Colin, 2015, p.25.

[22] La nécessaire évolution du droit des producteurs de bases de données pour permettre son adaptation à l'émergence du Big Data, Nicolas Courtier, Propriété Intellectuelle & la transformation numérique de l'économie, INPI, 2015.

[23] Ce dernier proposait, en complément d'une évolution de la législation, que l'INPI développe un outil électronique de déclaration des bases de données afin d'éviter toute insécurité juridique.

[24] *Australia Sup. Court, Telstra Corporation Ltd c/ Phone Directories Compagny Pty Ltd* (2010), FCAFC 149 § 335 ; *Hoge Raad*, 30 mai 2008, *Zonen Endstra c/ Nieuw Amsterdam* (cité *in* A. Lucas, H.-J. Lucas, et A. Lucas-Schloetter, *Traité de propriété littéraire et artistique*, LexisNexis, 2012, n° 57).

[25] "*L'intégration de l'intelligence artificielle dans l'ordre juridique : questions de temps*", Alexandra Bensamoun et Grégoire Loiseau, à paraitre.

[26] « La loi et le règlement peuvent comporter , pour un objet et une durée limitée, des dispositions à caractère expérimental »

[27] « La loi et le règlement peuvent comporter , pour un objet et une durée limitée, des dispositions à caractère expérimental »

Annexe 3

L'intelligence artificielle
et l'enseignement des technologies génériques

Contribution au groupe : Anticiper les impacts économiques et sociaux de l'intelligence artificielle

Auteurs:
Boris LE HIR (France Stratégie)
Olivier SAVRIMOUTOO (France Stratégie)

Le débat sur l'IA et l'emploi : une impression de déjà-vu

Les experts s'accordent à dire que l'intelligence artificielle représente un ensemble de techniques permettant à des machines d'accomplir des tâches et de résoudre des problèmes qui requièrent une certaine forme d'intelligence. Le Centre national de ressources textuelles et lexicales (CNRTL) définit l'intelligence artificielle comme « la recherche de moyens susceptibles de doter les systèmes informatiques de capacités intellectuelles comparables à celles des êtres humains ».

La baisse du coût de la puissance de calcul parallèle, l'accès à de grandes bases de données, ainsi que de nouveaux algorithmes sont autant de transitions qui ont permis aux technologies issues de l'IA d'accomplir des tâches et de résoudre des problèmes d'une complexité considérable. De fait, cette technologie représente un levier de compétitivité majeur pour les entreprises, qui y voient un moyen d'accroître leur productivité en automatisant certains éléments de leur processus de production.

Toutefois, les impacts économiques d'une adoption massive de l'intelligence artificielle dans plusieurs secteurs d'activités suscitent de nombreux débats.

D'un côté, la description de ces technologies suggère qu'elles ont pour vocation de remplacer l'humain dans certaines tâches, d'où les craintes sur l'avenir de l'emploi. Des travaux publiés par diverses institutions qui mentionnent qu'une part élevée de l'emploi serait affectée par ces technologies alimentent cette anxiété : Frey et Osborne[22] estimaient en 2013 que 47 % des emplois sont susceptibles d'être en partie automatisés aux États-Unis ; plus récemment, l'OCDE[23], le Comité d'orientation pour l'emploi[24] (COE) ou France Stratégie[25] ont évalué qu'une dizaine de pourcent des emplois, en moyenne, présentent un risque élevé d'automatisation, et

[22] Carl Benedict Frey et Michael A. Osborne (2017), "The Future of employment: how susceptible are jobs to computerization", *Technological Forecasting and Social Change.*

[23] Arntz, M., T. Gregory and U. Zierahn (2016), "The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis", OECD Social, Employment and Migration Working Papers, No. 189, OECD Publishing, Paris

[24] Conseil d'orientation pour l'emploi (2017), « Automatisation, numérisation et emploi, Tome 1 : Les impacts sur le volume, la structure et la localisation de l'emploi », Conseil d'orientation pour l'emploi, Paris.

[25] Nicolas Le Ru (2016) « L'effet de l'automatisation sur l'emploi : ce qu'on sait et ce qu'on ignore » *Note d'analyse*, France Stratégie.

que 50 % des tâches pourraient être automatisées pour 25 autres pourcent d'emplois ; enfin, le rapport de la Maison blanche sur l'automatisation affirme que ce sont les emplois les moins qualifiés qui encourent le plus grand risque d'automatisation[26].

D'un autre côté, les gains réalisés au travers de l'automatisation de certaines tâches pourraient être réinjectés dans l'économie par une baisse des prix, une hausse des salaires pour les employés (qualifiés) et/ou des profits plus importants pour les entreprises, ce qui aurait pour résultat une augmentation du PIB par habitant sur le long terme et une progression du niveau de vie.

Si la question d'un impact négatif sur l'emploi peut se poser, il reste que se limiter à une analyse d'effet direct est trop restrictif. Par le passé, des anticipations alarmantes sur l'emploi ont régulièrement été associées à des chocs technologiques, notamment dans l'agriculture et dans l'industrie, et se sont révélées en partie fausses. Elles étaient justement fondées sur une analyse limitée à l'impact direct sans prise en compte des autres effets induits potentiels et des effets de bouclage macroéconomique. Certes, à court terme, ces technologies sont censées accroître la productivité du travail et accélérer la substitution du capital au facteur travail. Certes, dans un monde où l'utilité marginale de la consommation est décroissante, l'accroissement de la demande issue de la baisse des coûts de production ne permettra sans doute pas de compenser cette moindre utilisation du travail pour une même quantité produite. Mais deux raisons principales font que cela ne suffit pas à affirmer que l'impact final sur l'emploi sera négatif. D'une part, il est faux de considérer que la composition de la production reste inchangée et que l'emploi perdu d'un côté ne peut pas être réalloué vers d'autres activités, éventuellement nouvelles. D'autre part, même si ces nouvelles technologies conduisent à rendre les processus de production plus efficaces et économes en travail, elles peuvent aussi contribuer à l'innovation de produit (apparition de nouveaux produits) permettant là encore de générer de la croissance et de l'emploi.

Les technologies génériques : un concept pour mieux penser l'impact de l'IA

Ces mécanismes sont mis en évidence par les théories économiques développées autour de la notion de technologie générique (TG). Celles-ci cherchent à déterminer les conditions sous lesquelles certaines technologies peuvent être des moteurs de croissance[27]. La littérature économique associe trois caractéristiques principales à ces technologies génériques[28] :

- Ce sont des technologies qui affichent une dynamique de développement soutenue ;
- Ce sont des technologies ubiquitaires, dans le sens où elles peuvent avoir de multiples utilisations dans de nombreux secteurs ;
- Ce sont des technologies dites habilitantes[29] qui offrent de nombreuses opportunités d'innovations additionnelles.

---

[26] Executive office of the President of the United States (2016), "Artificial intelligence, automation and the economy".

[27] Timothy Bresnahan et Manuel Trajtenberg (1995), « General purpose technologies, 'engine of growth'? » *Journal of Econometrics*.

[28] Ces propriétés sont souvent décomposées en quatre comme dans Lipsey Carlaw et Békar (2006), « Economic transformations ».

[29] Ce terme vient de l'expression : « enabling technologies ».

Ces propriétés sont alors capables de générer une spirale vertueuse entre les producteurs et les utilisateurs de la technologie. Bresnahan et Trajtenberg[30] montrent en effet que les technologies vérifiant ces propriétés sont à même de devenir des moteurs de croissance par l'existence d'interactions entre les producteurs et les utilisateurs. En particulier, ils décrivent deux types d'externalités qui engendrent une complémentarité entre les innovations des producteurs (les inventions) et celles liées à l'adoption de la technologie par les utilisateurs (les co-inventions), et entre les co-inventions elles-mêmes[31].

D'un côté, l'innovation du producteur de technologie générique génère des opportunités d'innovation chez les utilisateurs. Par exemple, le développement de l'IA pour le véhicule autonome peut conduire le secteur des transports à la création de nouveaux services de transport « intelligents à la carte ». En retour, les innovations chez les utilisateurs, qui augmentent les rendements d'utilisation de la technologie générique, vont accroître la demande pour ces technologies, donc les incitations à innover chez le producteur, les autres utilisateurs bénéficiant alors de ces nouvelles innovations induites. Si l'on reprend notre exemple du véhicule autonome, plus on aura inventé de façons d'utiliser le véhicule autonome, plus le marché sera important et plus il sera rentable pour les concepteurs de véhicules autonomes d'investir en recherche afin d'accroître leur performance, ce qui bénéficiera en retour aux utilisateurs.

Ces externalités spécifiques au concept de technologie générique impliquent que les rendements sociaux des investissements dans ces technologies sont supérieurs aux rendements privés, et cela de façon plus importante que dans le cas des innovations traditionnelles. En contrepartie, elles révèlent que des investissements complémentaires à ceux dans la technologie génériques sont nécessaires. Les utilisateurs doivent en effet investir (se former, se réorganiser, voire même faire de la recherche) pour être capables d'utiliser au mieux le potentiel de la technologie (en inventant de nouveaux usages).

Le déploiement des technologies génériques : un temps pour semer, un temps pour récolter[32]

Ces théories ont largement été développées et utilisées pour analyser l'impact économique de chocs technologiques historiques et en particulier pour expliquer les délais de leurs impacts. Paul David[33] par exemple rappelle que dans un premier temps la dynamo, inventée en 1900, n'a fait que remplacer la machine à vapeur sur les sites de production. Ce n'est que dans les années 1920, après l'invention du moteur unitaire (placé individuellement sur chaque machine), que les effets sur la productivité sont apparus. Grâce à cette « co-invention », les chaînes de production ont pu être réorganisées de façons beaucoup plus efficientes puisqu'il n'était plus nécessaire de disposer les machines le long d'un arbre de transmission alimenté par un moteur

[30] Ibid.

[31] Ces complémentarités caractérisent les secteurs producteurs et utilisateurs comme des compléments stratégiques (Bulow *et al.*, 1985).

[32] Helpman et Trajtenberg (1994), "A time to sow and time to reap: Growth based on general purpose technologies", National Bureau of Economic Research.

[33] Paul David (1991), "The computer and the dynamo: An historical perspective on the modern productivity paradox", The American Economic Review.

central (anciennement la machine à vapeur). La formalisation du concept de technologie générique a été développée au cours d'une période de doute sur le potentiel économique associé au progrès et au déploiement des technologies de l'information et de la communication (TIC) alors que ces technologies avaient suscité beaucoup d'attentes. Ce doute s'était notamment exprimé par le désormais célèbre paradoxe de Solow[34] : « Nous voyons les ordinateurs partout autour de nous, sauf dans les statistiques de productivité ». Face à ce constat, la formalisation du concept de technologie générique a alors mis en évidence que, si l'ordinateur est l'équivalent d'un vélo pour l'esprit[35], encore faut-il apprendre à pédaler. En d'autres termes, la diffusion des TIC ne suffit pas, il faut que des co-inventions par les utilisateurs se développent pour exploiter le potentiel de ces technologies. Ce n'est en effet qu'après 1995 que la croissance du PIB et de la productivité s'est accélérée aux États-Unis[36] et s'est propagée dans les secteurs utilisateurs intensifs où des investissements en formation et en capital organisationnel ont été nécessaires pour justement apprendre à utiliser ces nouveaux outils[37]. Cette approche permet donc d'expliquer en partie certaines dynamiques économiques mais elle éclaire aussi certaines divergences entre pays. Des études ont en effet montré que l'Europe a moins profité du développement des TIC que les États-Unis du fait à la fois d'un poids du secteur TIC plus faible et d'une moins bonne performance des secteurs de services intensifs en TIC[38]. Plusieurs facteurs (comme la taille du marché, la fragmentation linguistique, la régulation du marché du travail ou des biens et services par exemple) ont été avancés pour expliquer cette moindre capacité à développer et utiliser les TIC en Europe et ces analyses peuvent certainement apporter des enseignements pour le développement de l'IA en Europe.

L'Intelligence artificielle, une technologie générique ?

Un certain nombre d'éléments laisse penser qu'il serait pertinent de considérer l'intelligence artificielle sous l'angle d'une technologie générique pour étudier son impact sur l'économie et l'emploi. En premier lieu, il semble que ces dernières années le développement de l'IA bénéficie d'une forte accélération faisant suite : i) à la baisse du coût de la puissance de calcul parallèle ; ii) à l'accès à de grandes bases de données ; iii) à de nouveaux algorithmes tirant parti des deux développements précédents. Différentes études ont aussi mis en évidence le potentiel d'utilisation de l'IA dans de multiples secteurs, que ce soit dans la santé, le transport, la distribution ou dans l'industrie. Enfin, il paraît assez intuitif que le déploiement de l'IA ne se fera pas uniquement par des investissements dans la technologie mais nécessitera de lourds

[34] Robert Solow dans une declaration pour le New York Times Book Review, 12 juillet 1987.

[35] Steve Jobs: "What a computer is to me is the most remarkable tool that we have ever come up with. It's the equivalent of a bicycle for our minds." Memory and Imagination: New Pathways to the Library of Congress.

[36] Jorgenson, Ho et Stiroh, 2008 « A retrospective look at the US productivity growth resurgence », Journal of economic perspective; Cette, Mairesse, Kocoglu, 2009, "La productivité en France, au Japon, aux Etats-unis et au Royaume-uni au cours du XXème siècle" Revue de l'OFCE.

[37] Brynjolfsson et Saunders 2009, "Wired for information: how information technology is reshaping the economy" the MIT press.

[38] Van Ark, Inklaar, Guckin, 2008, "The productivity gap between Europe and the United-States: trends and causes", Journal of economic perspective; Dimelis et Popaioanou, 2010 « ICT growth effects at the industry level: a comparison between the US and the EU" Information economics and policy.

investissements complémentaires en formation, en réorganisation et très certainement en développement de nouveaux produits. Il est aussi très probable que cela conduise à repenser certains systèmes économiques dans leur ensemble, i.e. des changements de paradigme techno-économiques. Le cas de la voiture autonome fournit une illustration de tels changements potentiels : il impliquera sans aucun doute des changements majeurs dans les secteurs de l'assurance et des services de transports, y compris hors transports routiers comme le train qui pourrait perdre ou gagner en utilité selon que la voiture autonome deviendrait un complément ou un substitut à ce moyen de transport ; il remettra en question l'existence des écoles de conduite, modifiera le métier des garagistes et des assurances ; il fera émerger de nouveaux « business models » où les individus ne possèderont plus leur véhicule mais loueront un service de transport autonome à la carte avec un véhicule adapté à leur besoin à chaque déplacement, le tout avec un risque d'accident fortement réduit et sans problème de congestion. Il est donc tout à fait possible que l'IA entraîne l'innovation mais dans quelle proportion, la réponse reste bien entendu très incertaine.

Penser le changement de paradigme techno-économique

Sous cet angle, la question de l'impact sur l'emploi du déploiement de l'IA devient donc multiple:

- Quels emplois seront en effet directement touchés par le choc technologique ? Avec quelle ampleur et avec quelle dynamique ? En d'autres termes, quels sont les « métiers du passé » et quelle est la population la plus susceptible d'être impactée ?
- À l'inverse, quels sont les secteurs et les métiers vers lesquels les travailleurs déclassés par la technologie peuvent être redirigés ? En d'autres termes, quels sont les « métiers d'avenir » ?
- Quelles sont les possibilités de reconversion des « métiers du passé » vers les « métiers d'avenir » ? Sur ce point, la notion de compétence est bien sûr primordiale mais la dimension territoriale doit aussi avoir un rôle important dans l'analyse.
- Quels sont les secteurs faisant partie de la chaîne de production de l'intelligence artificielle et est-il important de les développer sur le territoire ? Le fait d'importer ou de produire la technologie est-il déterminant sur son effet ?
- Quels sont les secteurs potentiellement utilisateurs de l'IA et quel impact l'IA peut-elle avoir sur ces secteurs ? Va-t-elle uniquement conduire à accroître leur productivité ou va-t-elle aussi leur permettre d'innover et de développer de nouveaux biens et services ?
- Quels impacts potentiels globaux la technologie peut-elle avoir sur l'ensemble du système productif ? Quelle est l'ampleur du changement de paradigme techno-économique ?
- Quels sont les efforts à mettre en œuvre pour libérer le potentiel de croissance que ces technologies peuvent générer ?

Ce sont autant de questions auxquelles il est nécessaire de répondre pour déterminer si le développement et la diffusion de l'IA auront un impact négatif ou positif sur l'emploi et pour déterminer les meilleures stratégies à adopter en termes d'investissement dans la technologie, tant au niveau de sa production et de son utilisation, qu'en termes de formation des travailleurs pour anticiper le plus tôt possible les difficultés de transition.

Peut-on envisager des évaluations quantitatives e*x ante* ?

Une approche par analogie

Afin de prendre en compte les effets indirects potentiels du développement de l'IA, les auteurs d'une étude d'*Analysis Group*[39] ont tenté des approches par analogies.

Dans une première approche, Chen *et al.* considèrent que les principaux effets indirects du développement de l'IA pourraient être comparables, en proportion, au développement d'autres technologies. En particulier, ils s'intéressent aux externalités associées aux investissements en R & D privée et publique et en capital-risque. Ils reprennent pour cela les valeurs des rendements associés à ces investissements disponibles dans la littérature empirique[40] et déduisent que, selon ces valeurs, 1 $ investi en R & D privée, en R & D publique et en capital-risque dans l'IA rapporterait respectivement 1,99 $, 2,69 $ et 3,33 $ de PIB supplémentaire.

Cette approche est intéressante pour mettre en évidence les rendements sociaux supérieurs à l'unité des investissements dans la recherche et le développement d'une nouvelle technologie mais elle reste très approximative et, surtout, elle ne prend pas en compte les mécanismes économiques propres à ce type de technologie qui ont été décrits ci-dessus.

Dans une seconde approche, ces auteurs envisagent la possibilité que le déploiement de l'IA puisse avoir des effets similaires à ceux qu'ont pu avoir d'autres technologies par le passé. Dans cette approche, ils se fondent notamment sur les évaluations empiriques des impacts des technologies de l'information, de l'internet haut débit, de la téléphonie mobile ou encore de la robotique industrielle. En utilisant le développement et l'adoption de ces technologies comme *benchmark*, Chen *et al.* estiment des ordres de grandeur de l'impact que pourrait avoir l'IA sur l'économie. Si l'on admet que les technologies prises en compte ici sont génériques, cette approche est sans doute plus à même de prendre en compte les effets globaux de la technologie, à condition que les évaluations historiques intègrent bien ces mécanismes. Les résultats obtenus suivant cette approche sont nettement supérieurs à ceux issus de la première démarche. Néanmoins la transposition directe des impacts de ces technologiques à l'IA est difficilement envisageable ; il est certainement plus pertinent de chercher à décomposer les effets pour mieux appréhender les principaux canaux de transmission dans l'économie et de tester l'importance relative de ces canaux.

Le recours à la modélisation macro-sectorielle

Outre ces analogies directes qui peuvent être trop approximatives, il est possible d'approfondir le raisonnement par l'utilisation de modèles macro-sectoriels. Parmi ces modèles, Némésis[41]

---

[39] Nicholas Chen, Lau Christensen, Kevin Gallagher, Rosamond Mate, Greg Rafert(2015), "Global Economic Impacts Associated with Artificial Intelligence", Analysis Group Inc. financé par Facebook Inc.

[40] Ils se basent en particulier sur l'article d'Astrid Romain et Bruno Van Pottelsberghe (2004), « The Economic Impact of Venture Capital » qui évalue l'élasticité de l'output aux investissements en R & D privée, en R & D publique et en capital-risque à respectivement : 19,9 %, 13,6 % et 0,9 %.

[41] Le modèle Némésis est un modèle macro-sectoriel représentant 28 pays européens avec une décomposition sectorielle en 30 secteurs. Ce modèle a été développé par un consortium européen comprenant le laboratoire Erasme de l'Ecole Centrale Paris – relayé depuis par l'entreprise Seureco – le Bureau Fédéral du Plan belge et l'Université polytechnique nationale d'Athènes (ICCS/NTUA). Il a par ailleurs été en grande partie construit dans le

incorpore des mécanismes de croissance endogène et attribue aux TIC des propriétés de technologie générique. Il permet en ce sens d'apporter un certain éclairage à diverses questions sur l'impact de l'intelligence artificielle dans l'économie française.

Ces mécanismes sont introduits dans le modèle par une fonction de production d'innovation multidimensionnelle dépendant, dans chaque secteur, des investissements dans trois actifs : la R & D, les TIC et d'autres intangibles (logiciels et formation)[42]. En outre, le modèle distingue, dans les effets d'innovation, ceux d'augmentation de la productivité et ceux de création de nouveaux produits. Cette distinction est d'autant plus importante dans notre analyse que l'impact final sur l'emploi tient au fait que l'effet dominant du progrès technique est soit l'accroissement de la productivité et la moindre utilisation des facteurs, soit l'amélioration des produits et l'accroissement de la demande.

Toutes les précautions d'usage dans l'interprétation des résultats doivent bien entendu être prises puisque ce modèle ne représente pas explicitement le développement et la diffusion de l'IA. Néanmoins, le fait qu'il intègre des propriétés de technologies génériques associées aux TIC met en évidence les canaux de transmission des impacts de ce type de technologie.

La nécessité de tester la sensibilité des résultats aux différentes hypothèses

Comme nous l'avons décrit plus haut, l'analogie entre la diffusion de l'IA et celles des TIC de façon générale semble pertinente mais la transposition ne peut pas se faire directement. Sur la base de cette approche, il reste à déterminer quelles pourraient être les spécificités de l'IA par rapport aux logiciels classiques ou aux TIC en général. Pour cela, plusieurs questions se posent:

- Quels secteurs et quels pays sont les principaux producteurs d'IA?
- Quels secteurs sont potentiellement les principaux utilisateurs d'IA?
- L'IA a-t-elle des effets de complémentarité et d'externalité similaires à ceux des TIC et des logiciels?
- Les innovations induites par le déploiement de l'IA sont-elles davantage des innovations de produit ou des innovations de procédé comparées aux innovations traditionnelles moyennes? En d'autres termes, les paramètres de pondération entre les effets produits et les effets process dans le modèle doivent-ils être calibrés différemment?

Répondre à ces questions requiert un travail de futurologue dont les prospectives ne pourront être qu'incertaines.

Outre des analyses de scénarios simples, il est donc nécessaire, pour évaluer un champ des possibles, de réaliser des tests de sensibilité aux différentes hypothèses et aux différents calibrages du modèle. Des tests sur le poids des innovations de produits par rapport aux innovations de productivité, sur les complémentarités entre les facteurs d'innovations, sur les

---

cadre de projets de recherche européens et a bénéficié de l'apport de chercheurs de diverses institutions telles que le laboratoire MERIT de l'Université des Nations-Unies, le Centre for European Economic Research (ZEW) ou l'Université Libre de Bruxelles. Ce modèle a été jusqu'à présent principalement utilisé pour l'évaluation de politiques publiques au niveau national et européen dans les domaines de la fiscalité et du coût du travail, de l'énergie et de l'environnement et enfin dans le domaine de l'innovation.

[42] Cette extension des fonctions d'innovation du modèle a été réalisée grâce au développement des bases de données EU-KLEMS, qui fournissent des données sectorielles détaillées sur les investissements en TIC et en logiciel, et INTAN-Invest, fournissant des données pour divers investissements intangibles à des niveaux sectoriels. Ces bases de données, en particulier celles sur les actifs intangibles, sont encore en développement et pourront très certainement être étendues et améliorées.

externalités générées par les efforts d'innovation ou encore sur la part des importations dans les investissements en IA permettraient d'identifier l'influence de ces variables.

Conclusion

Évaluer *ex ante* l'impact économique que l'IA peut avoir dans les prochaines décennies (parce qu'il semble bien que ce soit l'horizon temporel nécessaire pour qu'une telle technologie déploie son potentiel) est pour ainsi dire mission impossible. Néanmoins, il est nécessaire d'anticiper suffisamment tôt les mutations pour en prévenir les effets négatifs et en saisir les opportunités. Pour cette raison, il est tout à fait utile de réaliser des exercices de prospective mais il est indispensable de considérer les changements dans leur complexité et de tenter de déterminer leurs principaux canaux de diffusion. Anticiper l'impact de l'IA au moins dans sa nature et, autant que possible, dans son ampleur ne peut donc se limiter à recenser les emplois que cette technologie pourra substituer mais doit prendre en compte des effets complexes caractéristiques de technologie dites génériques et des effets de bouclages macroéconomiques. Il est en effet utile d'identifier les métiers qui risquent de subir un déclassement ainsi que les risques sociaux associés au déploiement de l'IA mais il tout aussi primordiale de définir dans quels domaines et par quels moyens ces nouvelles technologies peuvent être un levier puissant. Les mutations que peuvent engendrer l'IA sont certainement importantes et il est très probable que celles-ci puissent être utiles dans la quête de nouveaux objectifs sociaux comme la transition écologique par exemple. Bien entendu, il ne s'agit pas de considérer que tout ce qui est nouveau est bon à prendre mais le fatalisme ou la crispation dans le changement n'aboutit en général qu'à en subir les effets négatifs sans les bénéfices. Il faut être acteur dans le changement pour être capable de prendre les bonnes décisions et de l'orienter de façon cohérente avec les préférences collectives.

Annexe 4 : membres du groupe de travail

Hindi	Rand	Président de Snips, membre du CNNum
Janin	Lionel	France Strategie
Julia	Charrié	France Statégie
Bonnet	Yann	CNNum
Berthet	Charly	CNNum
Cornut	Anne-Charlotte	CNNum
Levin	François	CNNum
Devillers	Laurence	Université Paris-Sorbonne, CNRS
Sylviane	Gastaldo	CGI
Laurent	Michel	CGI
Marion	Panfili	DGE
Babinet	Gilles	EU Digital Champion
Samuelian	Jean-Charles	PDG Alan.eu
Scherer	Eric	France Television
Nougué	Pierre	Ecosys Groupe et Reporters d'espoirs
Renard	Gregory	X Brain
Duan	Paul	Bayes Impact
Casili	Antonio	Télécom ParisTech
Dalle	Jean-Michel	Agoranov
Bensoussan	Alain	Lexing - Alain Bensoussan
Marie	Soulez	Lexing - Alain Bensoussan
Stiegler	Bernard	Institut de recherche et d'innovation
Boris	Le Hir	France Stratégie
Olivier	Savrimoutoo	France Stratégie

Annexe 5 : personnes auditionnées

Alexandre	Laurent	Co-fondateur de Doctissimo
Babinet	Gilles	Digital Champion
Benoilid	Daniel	FouleFactory
Bensoussan	Alain	Cabinet A. Bensoussan
Bertholon	Raphaëlle	CFE-CGC
Blanc	Pierre	Athling
Cadain	Alexandre	Hyperloop
Carrère-Gée	Marie-Claire	COE
Carron	Igor	LightOn
Casilli	Antonio	Telecom ParisTech
Chéron	Corentin	Airinov
Delvaux	Mady	Parlement européen
Drosne	Romain	RefundMyTicket
Duan	Paul	Bayes Impact

Dufoix	Diane	IBM
Escoubes	Franck	Bluenove
Georges	Bernard	Société générale
Goude	Yannig	EDF
Guellec	Dominique	OCDE
Guillaud	Hubert	InternetActu
Guillaume	Olivier	O² Quant
Humphrey	Éric	Spotify
Hwang	Tim	google
Kaplan	Frédéric	École Polytechnique Fédérale de Lausanne
Keller	Eric	FO
Legionnet	Gisèle	School of creative leadership
Luckin	Rose	UCL Knowledge Lab
Mardiné	Gérard	CFE-CGC
Massé	Alexis	CFDT
Molins	Jean-Luc	UGICT-CGT
Monier	Louis	Qwiki/Kiron
Mortensen	Dennis	X. ai
Moulier-Boutang	Yann	UTC Compiègne
Nougue	Pierre	Ecosys
Renard	Grégory	X Brain
Samuelian	Jean-Charles	Alan
Scherer	Éric	france TV
Stiegler	Bernard	Institut de recherche et d'innovation
Susskind	Daniel	Université d'Oxford
Thouvenel	Joseph	CFTC
Tolédano	Joëlle	CentraleSupélec
Turcq	Dominique	Boostzone
Vivenza	Jérôme	CGT
Volle	Michel	Institut de l'économie
Wolf	Clément	Google

GROUPE DE TRAVAIL

CARTOGRAPHIE DE L'IA

France IA - Groupe de travail Cartographie

Paysage des équipes académiques de recherche française contribuant aux avancées scientifiques de l'IA

1. Objectif de l'étude

Dans le périmètre de France IA, l'objectif du groupe de travail Cartographie était d'établir un premier état des lieux des compétences des équipes académiques françaises afin de contribuer à l 'élaboration et la mise en œuvre de la stratégie française.

2. Méthodologie

Sélection semi automatisée

Pour dresser un inventaire des équipes académiques Françaises dont le domaine de recherche peut être associé à France IA, les informations de présentations des équipes ont été croisées avec celles des projets en IA financés au plan national et européen. Nous n'avons donc pas retenu dans notre inventaire, les équipes « simple » utilisatrices de résultats en IA.

Ce travail a été réalisé par l'équipe France IA à partir de recherches sur le web et par l'exploitation des bases de données institutionnelles par l'outil ScanR.

Validation

A partir de ces éléments, un gros travail d'expertise, mené par le groupe de travail ad'hoc et renforcé par les interactions avec les autres groupes de travail, a permis d'identifier les principales équipes des domaines de l'IA. Sur cette base, et avec leurs données propres, la participation des tutelles a permis d'établir un inventaire validé des équipes et des compétences qu'elles développent.

3. Les premiers résultats

Les équipes identifiées par France IA

Nous avons identifié **268 équipes issues de 63 établissements ou organismes**. La listes des établissements, et des partenariats qu'ils ont établis, est présentée en annexe.

> *On observe que, presque systématiquement, une unité de recherche identifiée dans l'inventaire supporte plusieurs équipes, chacune traitant un champ spécifique des recherches en IA. Cette fragmentation, peut correspondre à une segmentation associée aux différentes thématiques, et s'explique couramment par une volonté de lisibilité et de reconnaissance dans une communauté scientifique précise. Ceci est donc dû à une des conditions de la performance scientifique, en l'état des processus d'évaluation. Cependant, la structuration d'une unité en équipes peut aussi être le reflet d'une stratégie d'organisation et la dispersion de chercheurs en IA dans plusieurs équipes d'une unité peut aussi être la conséquence du spectre large des thématiques couvertes par l'IA. Il serait donc abusif d'associer systématiquement un critère de qualité aux unités ayant plusieurs équipes identifiées dans notre inventaire.*
>
> ***Ainsi, le nombre d'équipes en IA d'une unité n'est qu'un indicateur, sans interprétation qualitative possible.***

A partir de cet inventaire, toutes ces équipes regroupent plus de **5000 chercheurs,** dont environ de 3000 sont en poste (les autres sont des doctorants, post-doc ou invités).

> *Il convient de noter que l'inventaire a permis d'identifier les équipes et, sur cette base, d'évaluer le nombre de personnes potentiellement liées aux recherches en IA. Cependant, dans une équipe, tous les membres ne sont pas systématiquement partie prenante d'une recherche directement sur une thématique liée à l'IA. Il suffit qu'un seul chercheur produise des résultats en IA pour que l'équipe soit identifiée.*
>
> ***Ainsi, le nombre de 5000 constitue une borne maximale de chercheurs en IA.***

Sur les 268 équipes identifiées**, 8 % relèvent des SHS.**

Pour affiner cet inventaire, il sera nécessaire d'utiliser des techniques de fouille de données plus puissantes.

Les tutelles

Pour avancer dans l'innovation et ses fondements, la recherche en IA s'appuie largement sur des organismes puisque **3/4 des équipes sont sous la tutelle de deux ou plus organismes.** Les tutelles des équipes, recensées par établissements sont principalement les universités (47), les écoles (13), le CNRS (58), INRIA (25), Le CEA (9), l'INSERM (3), l'INRA, IFSTTAR, IRD et l'ONERA.

Répartition géographique

La répartition géographique des équipes de recherche académique montre une adhérence aux lieux d'excellence identifiés dans France IA. Les différentes villes accueillant des recherches académiques en IA sont :

Angers, Avignon, Besançon, Bordeaux, Brest, Cachan, Caen, Calais, Chambéry, Clermont Ferrand, Compiègne, Evry, Grenoble, La Réunion, La Rochelle, Le Mans, Lens, Lille, Limoges, Lyon, Marseille, Metz, Montpellier, Nancy, Nantes, Nice, Nîmes, Orléans, Paris, Paris Saint Denis, Paris-Est, Pointe à Pitre, Poitiers, Reims, Rennes, Rouen, Saclay, Saint-Etienne, Strasbourg, Toulouse, Tours, Valence, Valenciennes

4. Conclusion

Le résultat obtenu est un premier niveau d'inventaire déjà largement très significatif pouvant être exploité pour orienter la stratégie France IA.

Afin de renforcer cet inventaire et exploiter plus avant les informations disponibles, le MENESR développera une cartographie sémantique automatisée que les parties prenantes pourront enrichir par un processus de *crowdsourcing* (avec validation).

Trois ensembles non disjoints d'équipes

Afin de faciliter l'inventaire des équipes, on pourra considérer trois secteurs qui, selon le niveau d'interdisciplinarité des équipes, ont des intersections.

- *STIC pour l'IA* : Equipes de recherche en IA dont les contributions relèvent des STICs
- *SHS pour l'IA* : Ceci concerne les équipes SHS développant des contributions pour l'IA (linguistique, sciences cognitives, sciences comportementales, etc.), mais aussi permettant d'étudier les conséquences de l'IA sur la société (éthique, droit, économie, sociologie, etc.).
- *IA pour les sciences*: Les équipes visent des résultats dans leurs domaines (Santé, Environnement, Transport, SHS, …) pour lesquels elles sont amenées à développer des contributions en IA.

Annexe : Liste des établissements et organismes identifiés par France IA

Etablissements et organismes	**Nombre d'équipes**
INRIA, CNRS, Université de Nantes, IMT	1
INRIA, CNRS, Université Joseph Fourier, UPMF , GrenobleINP, GIPSA Lab	1
Aix-Marseille Université	1
CEA	9
CentraleSupélec	2
CNAM	1
CNRS	10
CNRS - Ministère de la culture	1
CNRS-INSERM, Aix-Marseille Université	1
CNRS, Ecole Polytechnique	1
CNRS, Univ. Toulouse Jean Jaures	2
CNRS, Université de Caen, ENSI Caen	2
CNRS, Aix-Marseille Université	5
CNRS, Aix-Marseille Université, Université de Toulon, ENSAM	4
CNRS, Ecole centrale de Lille, Université de Lille 3	6
CNRS, Ecole Centrale Nantes, Ecole des Mines, Université de Nantes	5
CNRS, Ecole Normale Supérieure Lyon	2
CNRS, Ecole Normale Supérieure, Paris Saclay	1
CNRS, Ecole Normale Supérieure, Université Paris 3 Sorbonne Nouvelle	1
CNRS, Ecole Polytechnique	1
CNRS, EHESS, Aix-Marseille Université, Université Avignon	1
CNRS, EHESS, INRA, Université Toulouse 1	1
CNRS, ENS, EHESS	1
CNRS, IMT	3
CNRS, INSA, Université Lyon	5
CNRS, UGA, INPG	5
CNRS, Université Blaise Pascal, Clermont-Ferrand	4
CNRS, Université d'Artois	1
CNRS, Université d'Avignon	1
CNRS, Université de Bordeaux	3
CNRS, Université de Bretagne, Telecom	3
CNRS, Université de Caen	1
CNRS, Université de Caen, ENSI Caen	2
CNRS, Université de Cergy-Pontoise (UCP), École nationale supérieure de l'électronique et de ses applications de Cergy (ensea)	1
CNRS, Université de Cergy-Pontoise (UCP), École nationale supérieure de l'électronique et de ses applications de Cergy (ensea)	1
CNRS, Université de Franche Comté	2

CNRS, Universite de Limoges , Université de Poitiers	1
CNRS, Université de Lorraine	1
CNRS, Université de Montpellier	5
CNRS, Université de Nice	2
CNRS, Université de Toulouse	7
CNRS, Université de Tours	2
CNRS, Université de Valenciennes	1
CNRS, Université Grenoble Alpes	5
CNRS, Université Jean Monnet Saint Etienne	2
CNRS, Université Jean Monnet Saint Etienne, IMT	1
CNRS, Université Paris 1 Panthéon Sorbonne	2
CNRS, Université Paris Dauphine	1
CNRS, Université Paris Est Marne-la-Vallée, ENPC	1
CNRS, Université Paris Ouest	1
CNRS, Université Paris-Est Marne-la-Vallée	2
CNRS, Université Paris-Est Marne-la-Vallée, ESIEE Paris, École des Ponts ParisTech	1
CNRS, Université Paris-Sud	1
CNRS, Université Pierre et Marie Curie	12
CNRS, UTC Compiègne	3
Ecole Nationale d'Ingénieurs de Brest	1
Ecole Polytechnique, Mines ParisTech, Télécom ParisTech	1
ENS Ulm, INSERM	1
ENSTA	1
ESPCI	1
Grenoble INP	1
HEC	1
IBISC Evry	1
IFSTTAR	2
IMT	6
IMT Télécom Paristech	1
IMT, CNRS, Université de Bretagne Sud, Université de Bretagne Ouest, ENSTA, ENI Brest	2
INRA	2
INRIA	12
INRIA Université Paris-Dauphine	1
INRIA, CNRS, Institut polytechnique de Grenoble, Université Grenoble Alpes	2
INRIA, CEA	1
INRIA, CentraleSupélec	1
INRIA, CNRS, Université de Montpellier	2
INRIA, CNRS, Ecole centrale de Lille, Université de Lille 3	4

INRIA, CNRS, Ecole Polytechnique	1
INRIA, CNRS, ENS Ulm	1
INRIA, CNRS, Inserm, Université Pierre et Marie Curie	1
INRIA, CNRS, Université de Bordeaux, Bordeaux INP	1
INRIA, CNRS, Université de Lorraine	12
INRIA, CNRS, Université de Nantes	1
INRIA, CNRS, Université de Nice	1
INRIA, CNRS, Université de Rennes	1
INRIA, CNRS, Université de Rennes 1, ENS Rennes	1
INRIA, CNRS, Université Paris Sud	2
INRIA, IMT	1
INRIA, Université de Rennes 1, INSA	2
INSA Rouen, Université de Rouen, Université du Havre	1
IRD, Université Pierre et Marie Curie	1
IRT Jules Verne	1
ISAE Sup Aéro	1
National (CNRS, INRIA, etc)	1
National (opéré par CNRS et EHESS)	1
National (opéré par le CNRS)	2
OCDE	1
ONERA	1
Paristech, Ecole Polytechnique Mines ParisTech, Télécom ParisTech	1
Télécom Bretagne	1
Université Clermont Auvergne	1
Université d'Angers	3
Université d'Avignon	1
Université d'Orléans	1
Université d'Evry	1
Université de Caen	1
Université de Clermont Auvergne	1
Université de Guyane et des Antilles	1
Université de la Réunion	1
Université de la Rochelle	2
Université de Lorraine	1
Université de Lyon 1 - Claude Bernard	1
Université de Marseille	1
Université de Nîmes	1
Université de Poitiers	1
Université de Reims	2

Université de Savoie	1
Université de Valenciennes	1
Université du Littoral	3
Université Franche Comté	1
Université François-Rabelais	1
Université Lumière Lyon 2	1
Université Paris 1 Sorbonne	1
Université Paris 8	2
Université Paris Dauphine	1
Université Paris Est	1
Université Paris Lumières	1
Université Paris Sorbonne	2
Université Paris V	1
Université Paris-Est Créteil Val-de-Marne (Upec)	2
Université Rennes 2	2
Université Strasbourg	2

#FRANCEIA

FRANCE
INTELLIGENCE
ARTIFICIELLE

Cartographie IA

Paris, 21 mars 2017

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

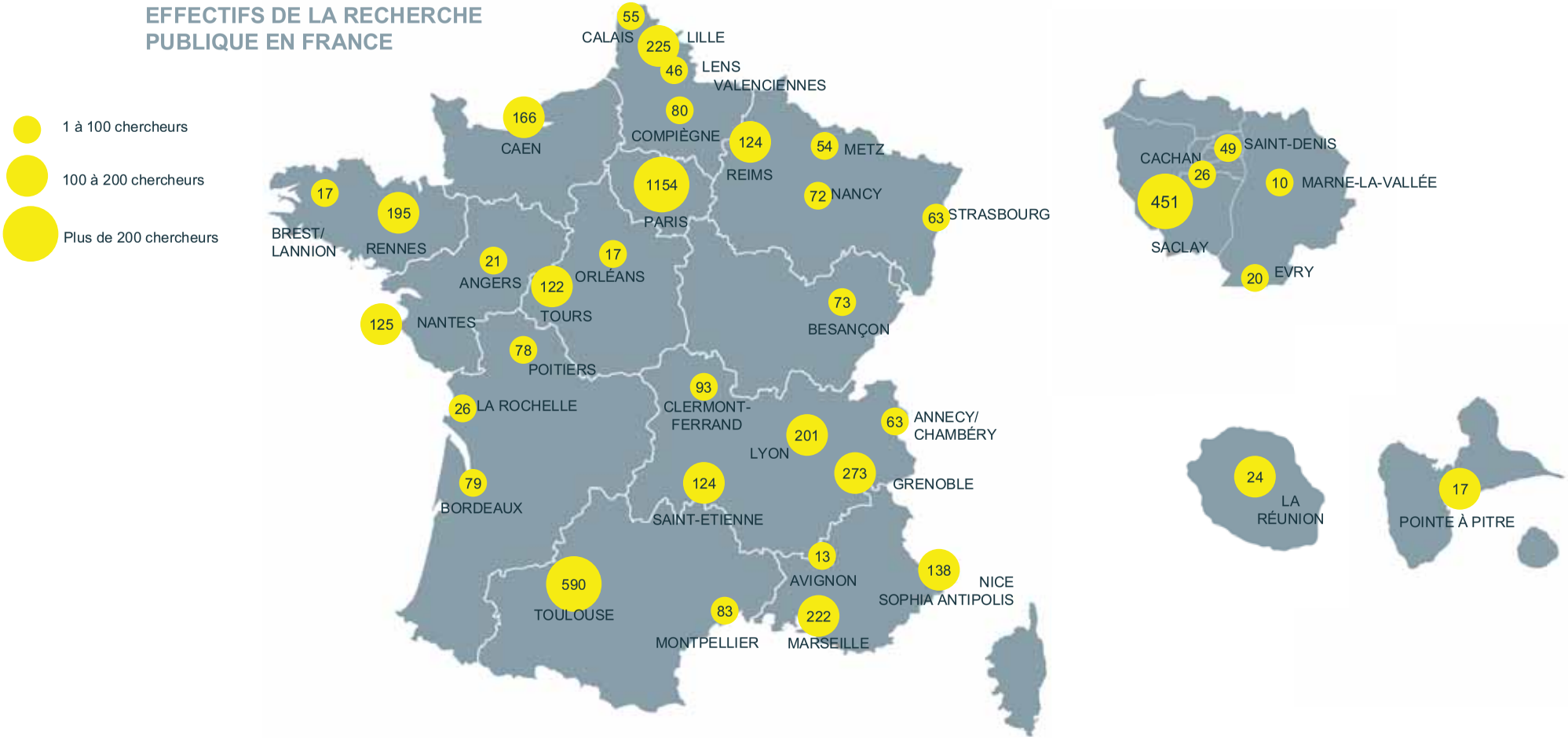
MINISTÈRE
DE L'ÉDUCATION NATIONALE,
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE

MINISTÈRE DE L'ÉCONOMIE
ET DES FINANCES

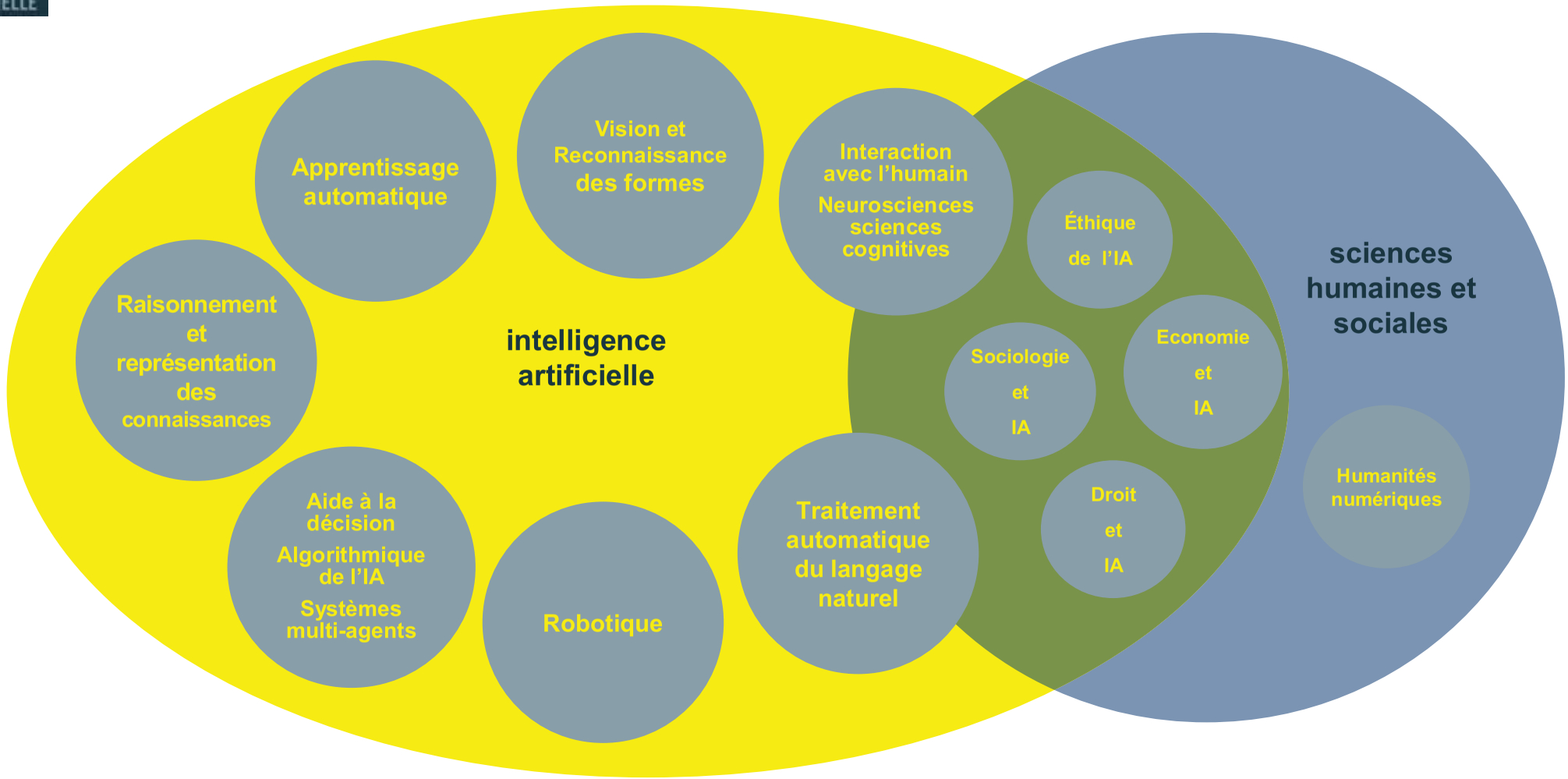
LES DONNÉES DE LA RECHERCHE

Environ 5 300 chercheurs, partout en France, dans 268 équipes identifiées (dont 8% relevant des SHS)

EFFECTIFS DE LA RECHERCHE PUBLIQUE EN FRANCE



Plusieurs domaines de recherche & domaines connexes SHS

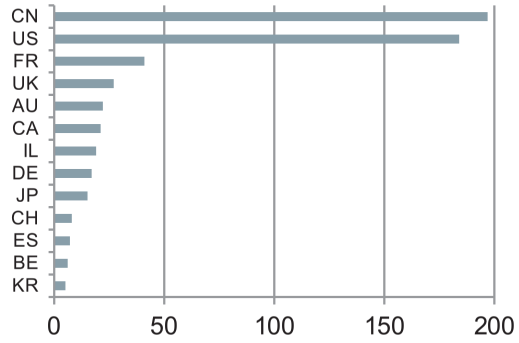


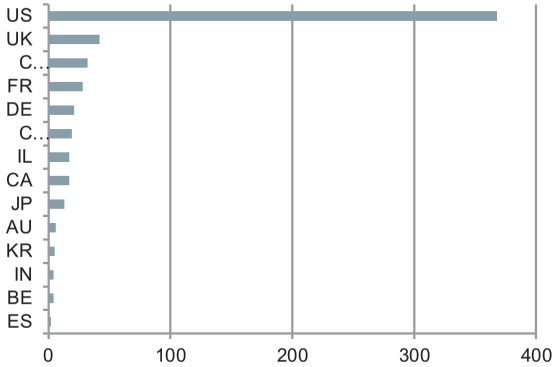
Une cinquantaine de sous-domaines de recherche en IA

IA et SHS	Représentation des connaissances	Apprentissage automatique	Traitement automatique des langues	Vision et Reconnaissance des formes	Robotique	Neurosciences, Sciences cognitives	Algorithmique de l'IA	Aide à la décision	Systèmes multi-agents	Interaction avec l'humain
Ethique de l'IA	Bases de connaissances	Apprentissage supervisé / non-supervisé / séquentiel et par renforcement	Analyse syntaxique Lexiques Discours	Parole	Conception	Compréhension et stimulation du cerveau et du système nerveux	Program-mation logique et ASP	Théorie des jeux	Coordination Multi-Agents (Planification multi-agents, Décision multi-agents)	Interaction avancée, apprentissage humain
Droit et IA	Extraction et nettoyage de connaissances	Optimisation	(Interaction, Connaissances et Langage Naturel)	Vision	Perception	Sciences cognitives	Déduction, preuve	Décision	Résolution Distribuée de Problèmes	Agents conversationnels
Economie et IA	Inférence	Méthodes bayésiennes	Traduction automatique	Reconnaissance d'objets	Décision		Théories SAT et Satisfaction de contraintes	Gestion de l'incertitude	Apprentissage multi-agents	Chatbots
Sociologie et IA	Web sémantique	Réseaux de neurones ou neuronaux		Reconnaissance d'activités	Action		Raison-nement causal, temporel, incertain	Explicabilité	Ingénierie Multi-Agents (Langages, plateformes, méthodologies)	
Humanités numériques	Ontologies	Méthodes à noyau		Recherche dans des banques d'images et de vidéos	Interactions avec les robots		Recherche heuristique		Simulation Multi-Agents (intéresse aussi les SHS)	
		Apprentissage profond		Reconstruction 3D et spatio-temporelle	Flottes de robots		Planification et ordonnan-cement			
		Fouille de données		Suivi d'objets et analyse des mouvements	Apprentissage des robots					
		Analyse de données massives		Localisation d'objets	Cognition pour la robotique et les systèmes					
				Asservissement visuel	Véhicules autonomes					

Benchmark international de la recherche
Généraliste: Chine #1, US #2, France #3
Spécialisées: US #1, GB/Chine #2/3, France #2/#4

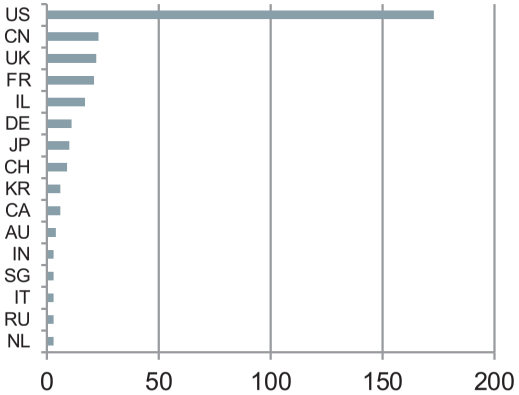
IJCAI 2016
 # sur 651 Articles

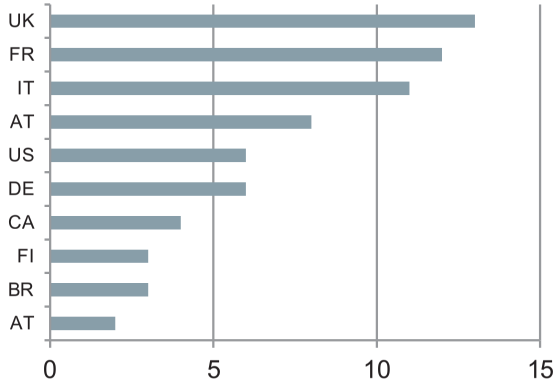




NIPS 2016
 # Articles sur 611

ICML 2016
 # Articles sur 324





KR 2016
 # Articles sur 68

(*) Source: Pages web officielles des conférences

Positionnement du volume d'articles publiées lors de conférences internationales en 2016, par des équipes de recherche françaises

FRANCE INTELLIGENCE ARTIFICIELLE

En 2016, parmi les 89 écoles d'ingénieurs et 45 universités disposant de formations liées à l'IA, on compte :

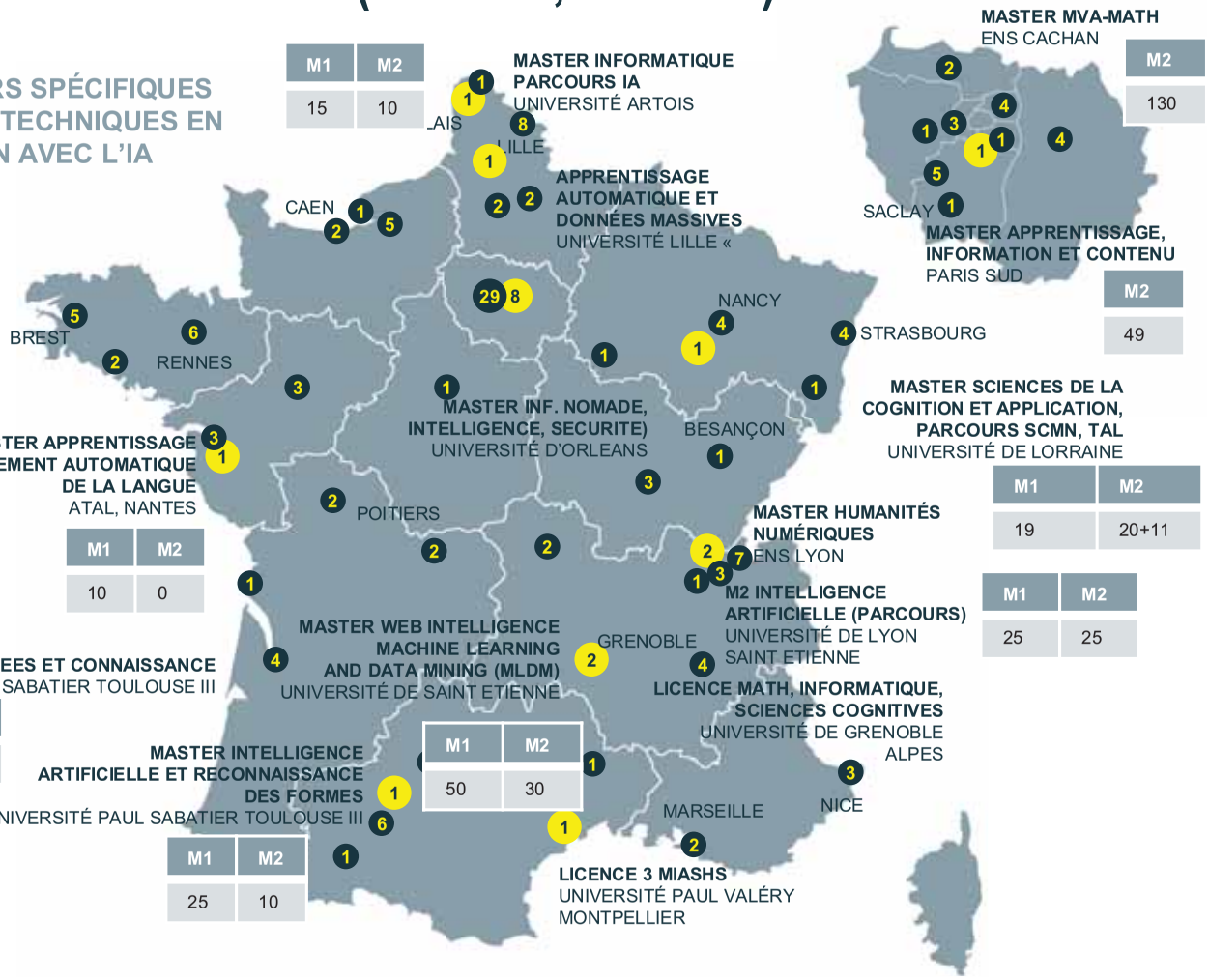
18 masters M1 et M2 spécialisés en IA, pour 1087 étudiants (M1: 415, M2: 672)

138 COURS SPÉCIFIQUES EN IA OU TECHNIQUES EN RELATION AVEC L'IA

M1	M2
15	10

1 Nombre de cours spécifiques IA ou en relation avec l'IA

1 Nombre de masters spécifiques IA ou en relation avec l'IA



M2
130

M2
49

M1	M2
19	20+11

M1	M2
25	25

M1	M2
10	0

M1	M2
20	20

M1	M2
50	30

M1	M2
25	10

DOCTORATS EN RELATION AVEC LES ÉQUIPES DE RECHERCHE ET ECOLES DOCTORALES

3 DOMINANTES

- APPRENTISSAGE, SCIENCE DES DONNÉES
- IA DETERMINISTE
- SCIENCES HUMAINES ET SOCIALES

8 MASTERS À PARIS

MASTER ANDROIDE: AGENTS DISTRIBUÉS, ROBOTIQUE, RECHERCHE OPÉRATIONNELLE, INTERACTION, DECISION
UNIVERSITÉ PIERRE ET MARIE CURIE PARIS 6

M1	M2
45	35

MASTER DAC : MASTER DONNÉES, APPRENTISSAGE ET CONNAISSANCES DE L'UPMC
UNIVERSITÉ PIERRE ET MARIE CURIE PARIS 6

M1	M2
34	34

MASH À DAUPHINE (ML APPLIQUÉ AUX SHS)
UNIVERSITÉ PARIS DAUPHINE

M1	M2
	23

MIAGE À DAUPHINE
UNIVERSITÉ PARIS DAUPHINE

M1	M2
75	134

INFORMATION, DE L'INTELLIGENCE AU CONTRÔLE
UNIVERSITÉ PARIS-EST

M1	M2
NR	NR

MASTER INFORMATIQUE EID2 : INFORMATIQUE DÉCISIONNELLE ET L'EXTRACTION DE CONNAISSANCES À PARTIR DE DONNÉES
PARIS 13 INSTITUT GALILÉE

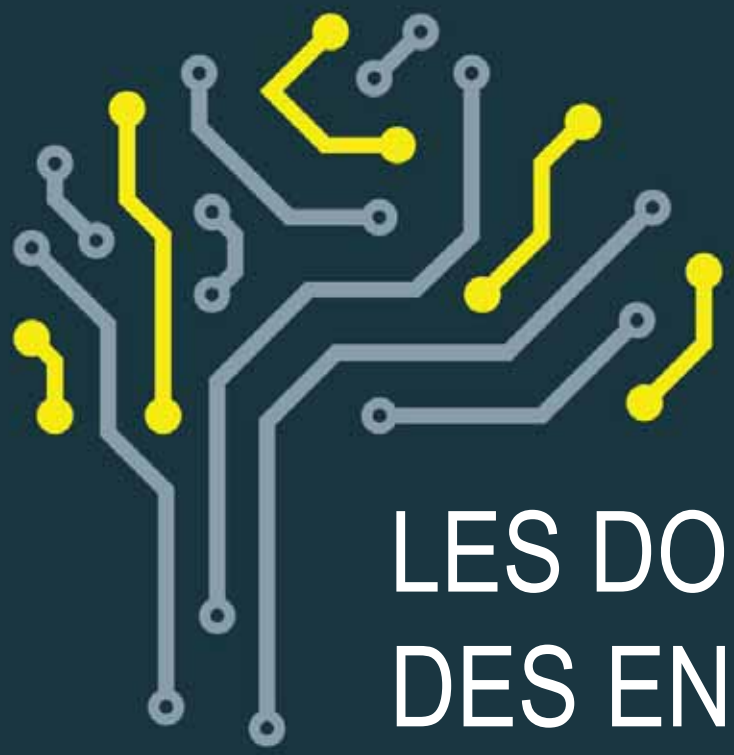
M1	M2
80	60

COG MASTER SCIENCES COGNITIVES
ENS ULM

M1	M2
NR	NR

MASTER INFORMATIQUE SPÉCIALITÉ IA UNIVERSITÉ
PARIS DESCARTES

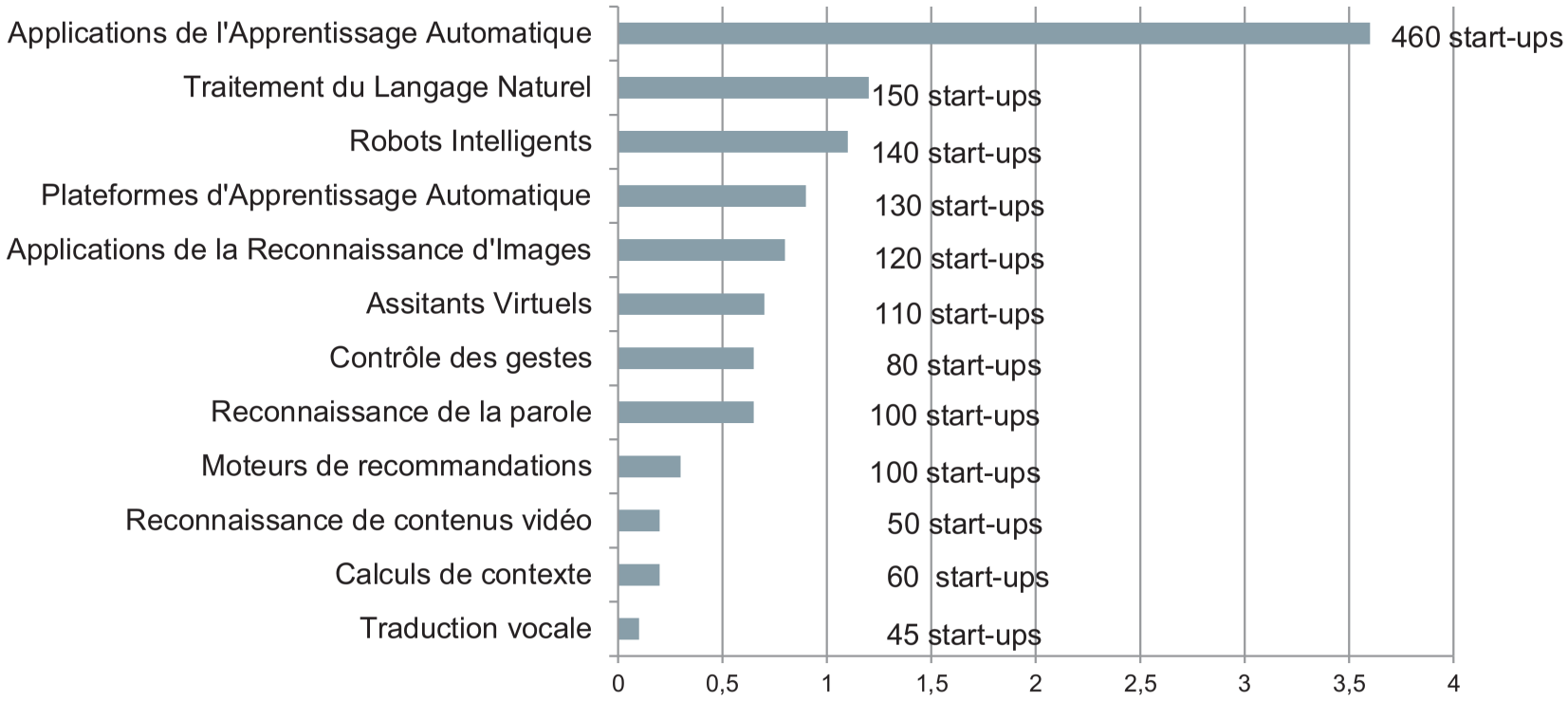
M1	M2
30	30



LES DONNÉES DES ENTREPRISES

70 pays : ~1500 start-ups en IA
~13 Md$ d'investissements cumulés 2012-2017

Distribution des investissements sur un échantillon de 1500 startups en IA dans le monde sur la période 2012-2017 (en Milliards de dollars)



(*) Source: Venture Scanner 2017
(**) Source: CB insights 2016

Etats-Unis: Consolidation de l'IA par acquisitions (*)
142 start-ups achetées de 2011 à 2017
42 start-ups achetées en 2016-2017, valeurs de $30m à $15Md

- Assistants virtuels
- Vidéo
- Prédictif client
- Cyber-sécurité
- Voiture autonome
- Santé

Deep learning
Agents conversationnels
Agents de voyages
Information contextuelle
Reconnaissance d'images

Volume acquisitions faible
Volume acquisitions modéré
Volume acquisitions important

Google (Alphabet)
Cleversense (2011)
DNNresearch(2013)
DeepMind (2014) $660m
Emu (2014)
Jetpac (2014)
Granata Decision Systems
Timeful (2015)
Api (2016)
Moodstocks (2016) 🇫🇷
Kifi (2016)
Dark Blue Labs (2014)
Vision Factory (2014)
Hark (2016)

Twitter
Madbits (2014) 🇫🇷
TellApart (2015)
Whetlab (2015)
Magic Pony (2016) $150m
Apprentissage et reconnaissance d'images.

Facebook
Face.com (2012) $55m
Wit.ai (2015)
Agents virtuels et recherche d'images

Microsoft
Netbreeze (2013)
Equivio (2015)
Genee(2016)
SwiftKey (2016) $250m
Maluuba (2017)
Assistants pour le droit et l'IA

IBM
Cognea (2014)
AlchemyAPI (2015)
Explorys (2015)
DeepL, santé, agents conversationnels (Watson)

Nokia (Finlande)
Desti (2014)
Medio Systems (2014)
Navigation (HERE maps)

Amazon
Orbeus (2015)
Angel (2016)
Harvest.ai (2017)
Apprentissage profond, reconnaissance images, assistants

Yahoo
Indisys (2013)
LookFlow (2013)
SkyPhrase (2013)

Oracle
Crosswise (2016) $50m
Palerra (2016)
Big data Client, Cyber-sécurité

Uber
Geometric Intelligence (2016)
Nouveau Lab.IA pour la conduite autonome

Intel
IQ Engines (2013)
Saffron (2015)
Itseez (2016)
Nervana (2016) $408m
Movidius (2016) $400m
MobilEye (2017) $15 000m

Samsung (Corée du Sud)
Viv Labs (2016)
Assistants virtuels.

Apple
Vocal IQ (2015)
Perceptio (2015)
Turi (2016) $200m
Tuplejump (2016)
Emotient (2016)

NICE (Israël)
Causata (2013)
Nexidia (2016) $135m
Fouille données client

Ford
SAIPS (2016) *Israël*
Argo.Ai (2017) $1Md
Apprentissage non-supervisé (vidéo) pour voiture autonome

General Electric
Wise.io (2016)
Bit Stew (2016) $135m
Apprentissage (Plateforme Predix), IoT

eBay
Hunch (2011)
Expertmaker (2016)
SalesPredict (2016)
Prediction achat client

SalesForce
Tempo AI (2015)
MetaMind (2016)
PredictionIO (2016)
Calendrier, Deep learning.

AOL
Gravity (2014) (*Recommandations*)
Convertro (2014)
Sociocast, Velos (2015)

🇫🇷 Start-up française

(*) Source: CB insights « Artificial Intelligence report » 2016
Les montants des acquisitions sont des estimations par des sites spécialisés

Stratégie des groupes internationaux (Etats-Unis)

	Recherche et Développement	Stratégie d'Acquisitions de start-ups en IA	Implantations R&D en France	Priorités pour l'IA
Google	Investissements en infrastructure et en recherche. Publications ouvertes de référence Partenariats en Europe avec des universités GB, Allemagne et Suisse. Investissements importants avec les universités decanadiennes autour du deep learning.	Poursuite active de leurs activités de croissance en IA par acquisitions de start-ups en IA.	Sans R&D en France Zurich (150 chercheurs), Londres, Dublin	Augmenter l'interactivité et connaître l'utilisateur pour proposer des services. Apprentissage profond, TALN, vision.
Facebook	Recherche active et présente sur le plan international (publications ouvertes de référence). Partenariats extérieurs sélectifs, notamment avec l'INRIA.	Peu de développement d'acquisitions en IA, à l'exception de Wit.ai en 2015.	Laboratoire de recherche en IA entre Paris (FAIR) et New York, avec une centaine de chercheurs.	Compréhension activités vidéo, traduction automatique. Apprentissage profond, TALN, vision, assistants.
SAP	Niveau de recherche interne en IA concentré sur quelques aspects du machine learning.	Besoins d'acquérir des compétences en IA, notamment par acquisitions.	Laboratoire en France, sur les usages de leurs produits (Levallois-Perret, Hauts de Seine).	« Machine learning » pour big data sur le cloud. Offrir des solutions au plus près du client.
Microsoft	Investissements massifs autour de la recherche l'IA et du cloud intelligent, en recherche et développement. Partenariats notamment en France avec l'INRIA (laboratoire commun). Investissements importants avec les universités decanadiennes autour du deep learning.	Développement affirmé de développement via l'acquisition de start-ups en IA. En particulier engagement avec le cluster Canadien de start-ups en deep learning.	Centres en IA autour de la région de New York (avec des plans pour la création d'un centre de 5000 ingénieurs/chercheurs en IA). En Europe le centre de M/S research à Cambridge.	Organiser le collaboratif H/M. Cloud intelligent (Azure), périphériques intelligents et assistants (Cortana, système à langage naturel).
Oracle	Très fort engagement dans le développement pour des nouveaux produits, mais avec moins de recherche qu'avec les groupes du GAFAM	Développement stratégique via des acquisitions en IA.	Centres de recherche autour des US: Californie et également Allemagne, Suisse, Australie. Certains développent des technologies en IA.	Développer des produits intelligents et adaptatifs, au plus près du client.
IBM	Engagement fort en développements autour de l'informatique cognitive et du système Watson, autour de l'apprentissage automatique, planification, représentation des connaissances, et les architectures cognitives.	Fort volume d'acquisitions en IA (activités autour de la santé). Développement interne ($1Md pour Watson, dont $100m pour acquérir des start-ups).	Le laboratoire d'IBM France concentre les recherches en région parisienne d'IBM. Les recherches plus significatives sont effectués dans leurs centres de recherche aux Etats-Unis.	Transformer leurs produits en « informatique cognitive », proposer des assistants (Watson=2000 employés). TALN et apprentissage renforcé.

Investissement en recherche en IA important
Investissement en recherche en IA moyen
Investissement en recherche en IA faible

Volume d'acquisitions en IA important
Volume d'acquisitions en IA moyen
Volume d'acquisitions faible

(*) Source: Auditions 8 mars avec groupes internationaux
(**) TALN: Traitement Automatique de Langues Naturelles

Stratégie des Groupes Internationaux (Asie)

	Recherche et Développement	Stratégie d'Acquisitions de start-ups en IA	Implantations R&D en France	Priorités pour l'IA
Baidu	Big data lab (Pékin) investissement de 300m€. R&D=14% du CA=1,4Md Institute of deep learning (Pékin)	Raventech (Chine, 2016), Assistant Voix 2016: Création de fonds De $200m pour investir en IA et RA.	Silicon Valley AI lab à San Francisco, Etats-Unis.	Apprentissage profond, TALN (deep speech)
Tencent	Création en 2016 du Tencent AI lab avec 30 chercheurs permanents.	Acquisition de « Imagu Vision tech » en 2016, entreprise basée en Israel.	Laboratoire de recherche à Shenzhen, Chine.	Machine learning pour recevoir des articles de presse spécialisés. WeChat, QQ, jeux vidéos.
Xiaomi	Création d'un laboratoire de recherche spécifique à l'IA en 2016.	Sans stratégie spécifique d'acquisitions de start-ups en IA	Laboratoire de recherche 'Xiaomi Exploration Lab » à Pékin, Chine. Laboratoire dédié aux projets avancés, en IA en particulier.	Volonté d'adopter IA dans tous leurs services (smartphones). IoT et IA dans tous les services et VR
Alibaba	Création d'un laboratoire spécifique dédié aux recherches en IA	Sans stratégie spécifique d'acquisitions de start-ups en IA	Laboratoire de recherche à Pékin, Chine.	Cloud intelligent Plateforme IA pour concurrencer Amazon
Huawei	Recherche intensive en IA. Partenariat de recherche en IA avec l'université de Californie à Bekeley en 2016 pour 1 millions de dollars en TALN (**)	Sans stratégie spécifique d'acquisitions de start-ups en IA	16 centres dans le monde. En France centres (Sophia-Antipolis, Boulogne-Billancourt). Centres autour de l'internet des objets, traitement d'images, design et mathématiques.	Développement de Smartphones dotés d'IA, avec des fonctions d'apprentissage automatique.
Samsung	Membre stratégie nationale IA Coréenne	Viv Labs (2016) Assistant virtuel. Augmentation des acquisitions à venir	Equipes de recherche et développement implantés en Russie, Inde, Israel, Etats-Unis, Chine Pologne	Assistants virtuels (coopération avec Google sur l'IA).
Fujitsu	Centre de recherche IA autour du deep learning, multimedia, connaissances. Fujitsu est leader en PI en IA.	Sans stratégie spécifique d'acquisitions de start-ups en IA	Centre dédié aux technologies du numérique en France (2017), représentant 50m€ d'investissement avec l'Ecole Polytechnique.	Human centric AI – Zinrai TALN, reconnaissance d'image, planification, prédiction.
Panasonic	Centre de recherche: machine learning, sensing, actuation. Panasonic-AIST Advanced AI Cooperative Research Laboratory	Fonds $10m pour acquisitions/JV pour smartphones	Présence d'équipes de recherche en Silicon Valley et Singapour.	Traduction automatique, coopération robot / humain, détection de piéton haute précision

(*) Source: Pages officielles des entreprises
(**) TALN: Traitement automatique du langage naturel.

Investissement en recherche en IA important
Investissement en recherche en IA moyen
Investissement en recherche en IA faible

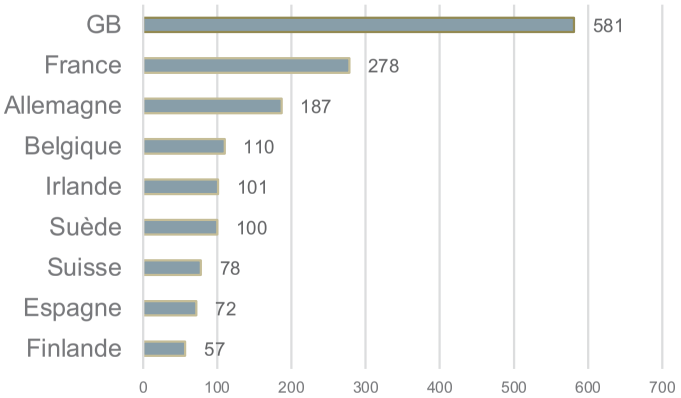
Volume d'acquisitions en IA important
Volume d'acquisitions en IA moyen
Volume d'acquisitions faible

Actions gouvernementales: « Benchmark » international

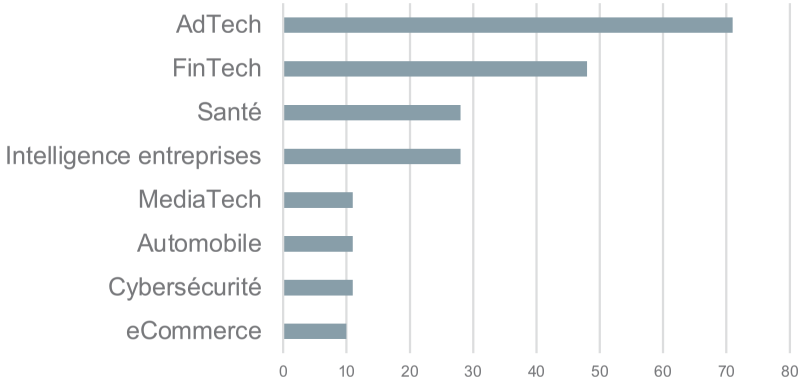
	Plans nationaux concernant l'IA Actions sociétales autour de l'IA	Recherche Publique: institutions, projets, investissements	Formation(s) en IA	Soutien aux entreprises dans l'adoption de l'IA
Etats-Unis	2016 Maison Blanche : « Future of Artificial Intelligence" et "National AI R&D Strategic plan" et "AI, Automation, and the Economy" 2014: rapport Big Data: Seizing opportunities, Preserving Values	$1,1Md en 2016 investis en Recherches liées à l'IA. Demande budget DARPA 2017 en IA: ~$12Md. Budget I-ARPA: confidentiel.	50 universités avec programmes spécifiques en IA.	La recherche privée (grands groupes) transfère les résultats de la recherche publique avec la mobilité des chercheurs et professeurs.
Canada	Sans plan spécifique national.	Investissements centres de recherche canadiens de l'ordre de 200m€	Excellence mondiale autour des universités d'Alberta, Toronto, et Montréal.	Sans actions spécifique.
Chine	Plan quinquennal 2016-2020 : Priorité big data et production intelligente. Plan 2030 : Accélérer l'usage de l'IA. Plan made in China 2025 : accélérer l'automatisation : 70% robots, 30% systèmes intelligents et placer 2 entreprises chinoises au Top 5 mondial en robotique. Plan Internet+ (2017): fonds de 13Md€ dédié au numérique et la construction d'une filière IA.	Plan Internet+ plateforme gouvernementale de soutien à l'IA: fonds VC en IA de 140m€ géré par la CAS. Investissements publics chinois 2015 en IA en recherche: 185m€ « China Brain project » lancé en 2016 (fonctions cognitives du cerveau)	Nombreuses universités chinoises dispensent des cours et des programmes sur l'intelligence artificielle, plutôt de niveau master.	Via le plan Internet+ Difficultés du dispositif chinois de transfert public/privé.
Japon	Super Smart Society 5.0 plan 2015: New Robot Strategy lancée par le METI 2016 : Lancement AI Center dans le cadre du Programme « next-generation AI technology » lancé par le MEXT. AIST centre de coopération et R&D sur l'IA (AIRC)	Augmentation budget de recherche en IA 250% de 2016 à 2017. 2015: Advanced 2015 Integrated Intelligence 57m€ AIRC: Platform Project: dont 3000 membres: 157m€	Les formations en IA et techniques associés sont réparties sur un groupe spécifique d'universités.	Exemple d'initiative: consortiums public/privé de 50 entreprises pour développer les usages de l'IA dans le médical.
Corée du Sud	Centre R&D en IA = 800m€ lancé en 2016 Mai 2016 doc stratégique sur les sciences du cerveau. 2016: Stratégie des 7 grands industriels Coréens sur l'IA	Budget recherche IA 118 M USD sur un budget national R&D de 16,6 Md USD.	Formations sur quelques universités.	MSIP (IITP) favorisent le transfert technologique entre instituts/universités et les PME.
Allemagne	2013: Plan Industrie 4.0 2017 : Plan Cyber valley autour du deeplearning	DFKi centre de recherche en IA 870 employés, 680m€. Centre Max Planck. 200 thèses en IA en 2016	22 universités proposent des cursus spécifiques en IA. 80 chaires en IA en Allemagne.	Politique de transfert via la coopération des centres de recherche et doctorands.
Royaume-Uni	IA fait partie du Digital Plan 2017 (dont budget pour l'IA £17 milions)	2015: création de l'Alan Turing Institute : 45m£ / 5ans	20 universités proposent des cursus spécifiques en IA, et une croissance en doctorats et chaire data science.	Actions spécifiques autour du digital plan 2017

Start-Ups en IA en Europe, 22 pays
2016: 1,8Md€ d'investissements, 306 transactions

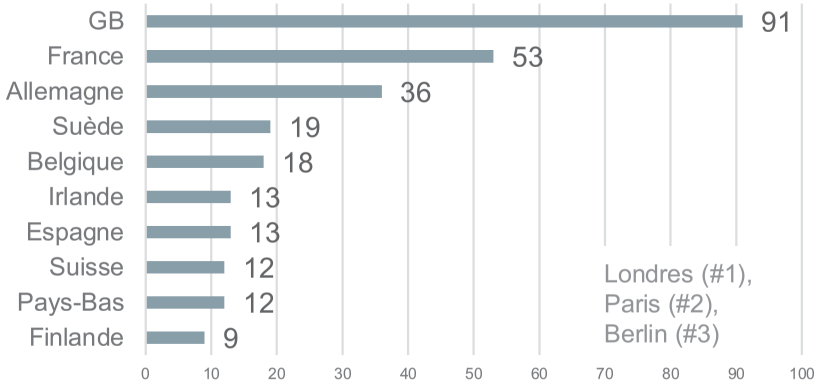
2016 (M€): Levées de fonds par pays



2016: Nombre de transactions par filière en Europe



2016: Nombre de transactions par pays en Europe



Source: (2016 European AI Scale-Ups; Sirris, Omar Mohout)

Investissements en IA en Europe par des grands groupes

1. Salesforce
2. RobertBosch
3. Intel
4. Alstom
5. Publicis
6. Samsung
7. AxelSpringer
8. SchneiderElectric
9. SolvayRhodia
10. SNCF
11. Mitsui
12. Orange
13. Nokia
14. Grupa Pracuj

Investisseurs les plus actifs en Europe:

1.KimaVentures (France)
2. IBB (Germany)
3. Salesforce Ventures (USA)
4. Ventech (France)
5. Sunstone Capital (Sweden)
6. Notion Capital (UK)
7. IQ Capital (UK)
8. Amadeus Capital Partners (UK)
9. BPI France (France)

France: Filières industrielles : développements et partenariats autour de l'IA

FILIÈRES INDUSTRIELLES ET PRIORITÉS IA	ENTREPRISES
Aéronautique, Espace Développement de systèmes autonomes (avions, imagerie satellitaire), et leur circulation dans l'espace aérien et en orbite Automatisation de la production et de la maintenance des plateformes et systèmes. Conception de produits complexes assistée pour les bureaux d'études avec aide à la décision.	***Airbus Group*** Avions et hélicoptères à la demande. Airbus Smarter Fleet (avec IBM). Production: Robots humanoïdes (projet FUTURASSY avec Kawada). Inspection automatisée. Détection d'images satellitaires (avec Google, projet OneAtlas). ***Thales***: Cybersécurité. Circulation aérienne intelligente. SAFRAN Aide à la décision ***Dassault*** Systèmes de pilotage autonome pour avion de combat.
Banque, Assurance Recherche et analyse des clients approfondie (assistance conseillers). Systèmes anti-fraude	***Crédit mutuel***: partenariat IBM Watson. ***BNPP***: investissement dans Smartly.ai (Chatbots). Partenariat avec Critéo. ***SG***: assistants conseillers. ***Axa***: création d'un fonds pour la recherche finance des chercheurs qui travaillent sur l'IA et sur la confiance dans le Big Data
Energie, Environnement Surveillance de sites industriels. Villes intelligentes: exploitation des données clients avec capteurs intelligents (smart grid et IoT)	***eDF***: développement de réseaux intelligents (smart grids). ***Veolia***: récolte et tri de déchets intelligente (avec Huawei)
Distribution, luxe, tourisme (B2C) Aide à la recherche produit/service et à la décision des clients Placement produit multi-canal optimisés.	***L'Oréal***: investissement en systèmes prédictifs et relation client. ***Accor***: intégration moteurs IA, pour CRM et marketing avancé. ***Publicis***: système Cyc pour aide à la décision.
Santé Développement de traitements plus efficaces. Exploitation de données santé clients adaptées.	***Sanofi***: co-entreprise Onduo (avec Google) lutte contre la diabète. Traitement du cancer avec IBM (Watson). Accès à une base de données de 118m de patients FDA (US). ***Dassault Systèmes*** Optimisation processus de R&D (avec la société Ipsen dans le cadre du consortium BioIntelligence).
Transports Analyse clients. Développement de systèmes autonomes.	***Alstom***: investissement navettes autonomes EasyMile. ***Renault/Valéo***: développement de systèmes de conduite autonome. Investissement dans des capteurs intelligents (avec Chronocam). ***PSA***: prototype de voiture autonome Picasso C4.
Télécoms et electronique Gestion et optimisation du réseau. Développement et traitement automatique de la relation client. Interfaces client (chatbots, box en connexion avec l'environnement).	***Orange***: apprentissage automatique appliqué à la relation clients et centres d'appel. ***Bouygues***: monétisation des données et systèmes anti-fraude. ***Nokia***: augmentation des services de localisation, réseaux programmés et automatisés. ***Technicolor:*** applications vidéo. ***ST Microelectronics*** : contrôle qualité en production. Capteurs IoT. ***Gemalto***

- ***Autonomie***
- ***Cyber-sécurité***
- ***Prédictif client***
- ***Santé***
- ***Assistants***

(*) Source: Pages officielles des entreprises

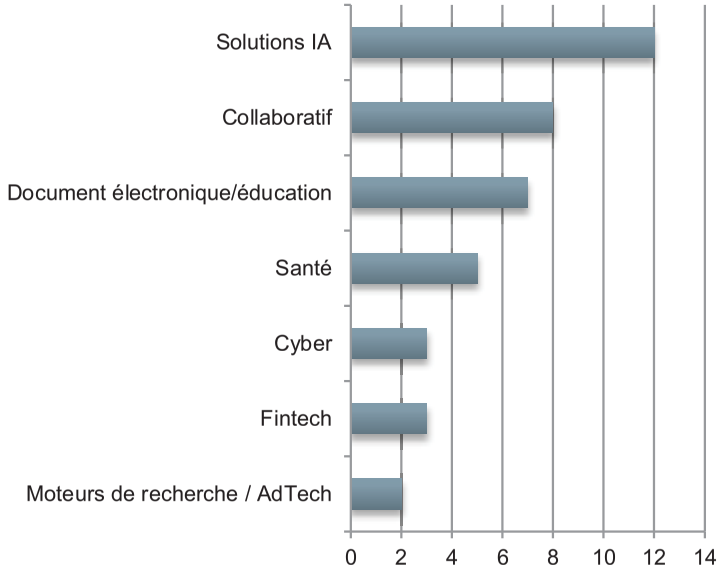
80 ETI et PME de l'IA en France: 40 en création d'IA et 40 faisant usage de développements en IA
Montée de l'usage du machine learning et du traitement du langage naturel

Utilisateurs d'IA

Apprentissage Automatique
Atos-Bull
Criteo
Vente Privée
Withings
Showroom Privé
SeLoger
Kelkoo
PriceMinister (Rakuten)
Meetic

Développements en IA

Santé	Soutions IA	Collaboratif	Document électronique/ Education
ARCAN: Santé, médico-social MEDASYS : Santé H2AD Carestream: Santé, imagerie médicale Synovo : transport sanitaire	**Atos-Bull** **MASA Group** AMI Software: prise de décision ARDANS: gestion des connaissances Teradata: Big Data UPETEC: valorisation des données MONDECA: modélisation des données KLS-Optim: Optimisation aide à la décision Eolas (Atos): Big Data Eurodecision: recherche opérationnelle et décision NOOPSIS : fouille données SYNCHRONEXT: web sémantique	Jamespot PERTIMM: pertinance recommandation client PI3C: ingénierie colaborative Reviatech: VR/AR Interaction games: serious games Vigiglobe: réseaux sociaux VIACCESS-ORCA VISEO	IMMANENS Sejer Openclassrooms MAXICOURS : éducation MANTANO : distribution d'e-books Exomakina Gayatech
Moteurs de recherche AdTech		**Fintech**	**Sécurité**
Moteurs de recherche AdTech		Finaxys: IT Banque finance assurance KYRYBA : Fintech TEMIS: sûreté fiduciaire	I-TRUST : Cybersécurité et machine learning Intrinsec Wallix Nexyad Chronocam



(*) Source: Pages officielles des entreprises, AFIA.

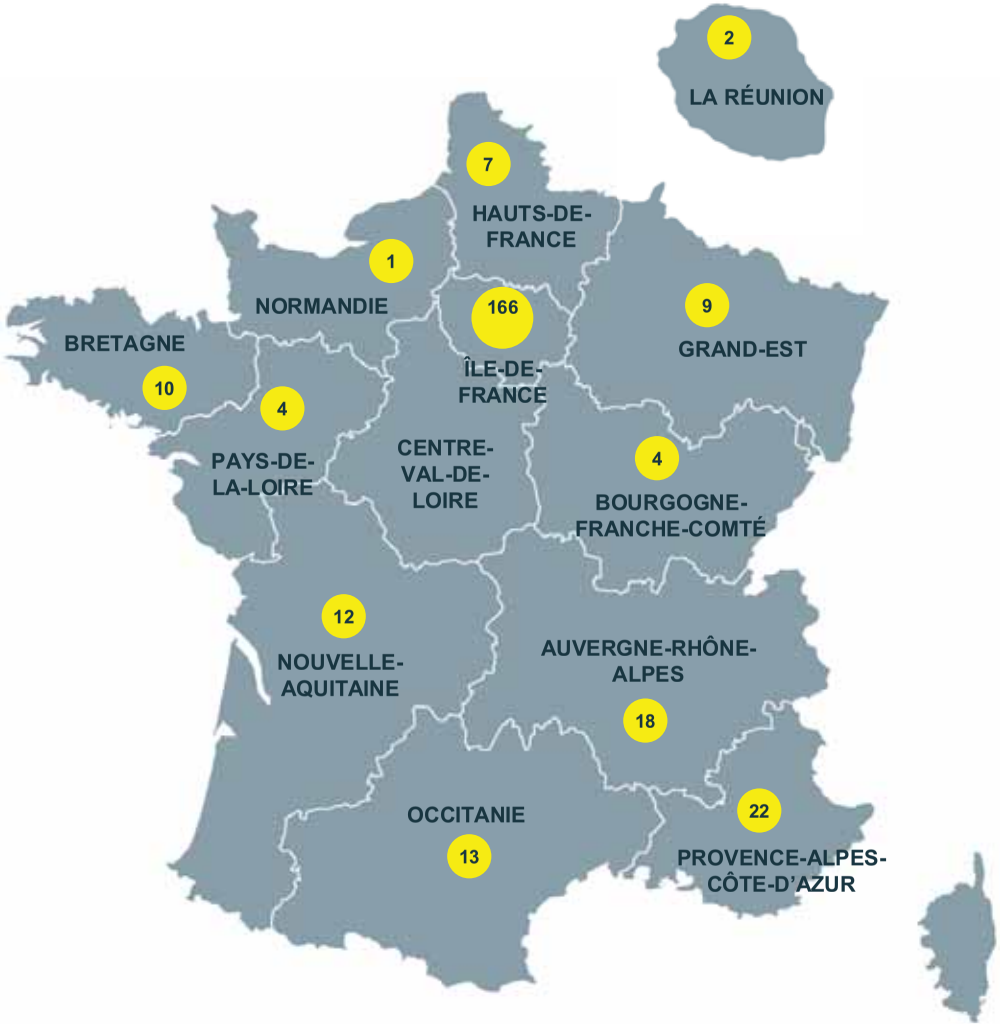
Plus de 270 start-ups en France spécialisées en IA

START-UPS FRANÇAISES INSTALLÉES À L'ÉTRANGER

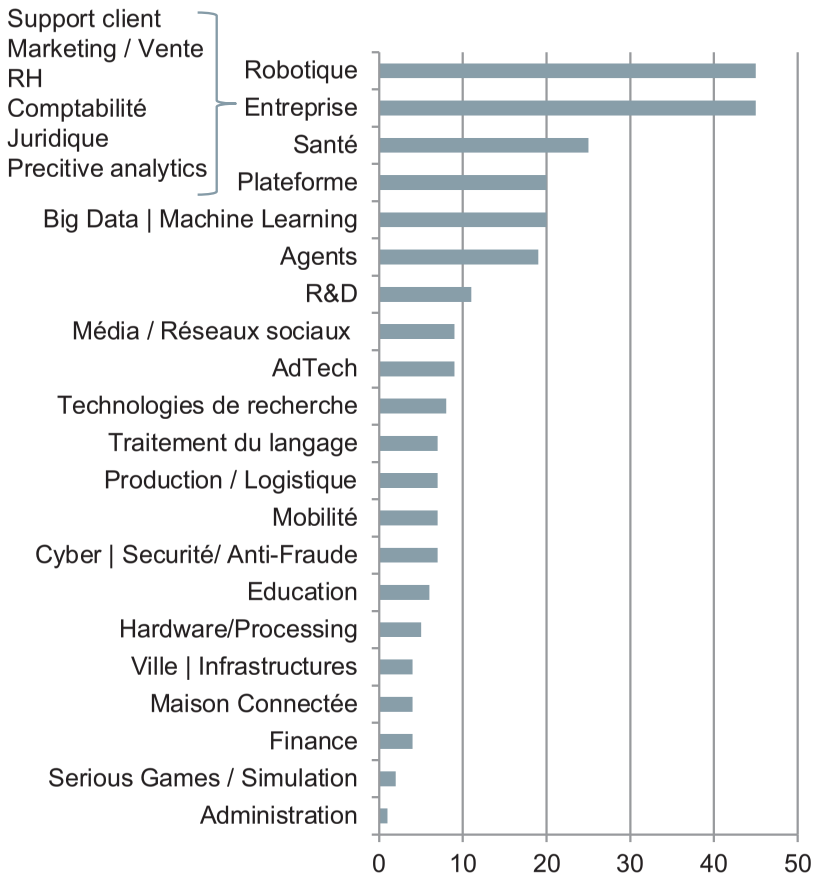
ÉTATS-UNIS
BAYES
IMPACT
YSEOP
PRESTASHOP
ADOMIK
GORGIAS
OWKIN
SEARCH'XPR
PLACEMETER
OTOSENSE
WIT.AI
WCA
ROBOTICS
RHYTHM
INTUITIVE
SURGICAL
LORE.AI
OTOSENSE

ROYAUME-UNI
BIGROBOTS

RUSSIE
BRAINIFY



Créations de start-ups 2000-2016 : 270
Services aux Entreprises, Robotique et Santé en tête
Agents et « Machine learning » en augmentation



RYTHME DE CRÉATION DE START-UPS EN IA