

ÉTUDES ÉCONOMIQUES

PROSPECTIVE

Intelligence artificielle :
État de l'art et perspectives pour la France

Synthèse

Date de parution : février 2019
Couverture : Hélène Allias-Denis, Brigitte Baroin
Édition : Martine Automme, Nicole Merle-Lamoot

ISBN : 978-2-11-152634-1
ISSN : 2491-0058

Intelligence artificielle : État de l'art et perspectives pour la France



Synthèse

Le Pôle interministériel de Prospective et d'Anticipation des Mutations économiques (Pipame) a été créé en 2005. Son objectif est d'éclairer, à un horizon de cinq à dix ans, les mutations qui transforment les principaux secteurs économiques.

La mondialisation transforme les entreprises et les pousse à s'adapter à une concurrence accrue sur des marchés plus diversifiés et lointains. Le numérique, avec l'essor de la robotique, de l'intelligence artificielle, des objets connectés, etc. transforme les entreprises tant dans le secteur de l'industrie que dans celui des services.

Dans ce contexte, le PIPAME apporte aux acteurs publics et privés des éléments d'alerte et de compréhension de ces mutations. Il propose des préconisations d'actions à court, moyen et long terme, afin d'accroître la compétitivité des entreprises françaises. Le Pôle aide les professionnels et les pouvoirs publics dans leur prise de décision.

Le secrétariat général du Pipame est assuré par la sous-direction de la Prospective, des Études et de l'Évaluation Économiques (P3E) de la direction générale des Entreprises (DGE).

Les départements ministériels participant au Pipame sont :

- le ministère de l'Agriculture et de l'Alimentation,
- le ministère des Armées,
- le ministère de la Cohésion des territoires et des Relations avec les Collectivités territoriales,
- le ministère de la Culture,
- le ministère de l'Économie et des Finances,
- le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation,
- le ministère de l'Europe et des Affaires étrangères,
- le ministère de l'Intérieur,
- le ministère des Solidarités et de la Santé,
- le ministère des Sports,
- le ministère de la Transition écologique et solidaire,
- le ministère du Travail.

Avertissement

La méthodologie utilisée dans cette étude ainsi que les résultats obtenus relèvent de la seule responsabilité du prestataire ayant réalisé cette étude, *Atawao Consulting*. Ils n'engagent ni le PIPAME, ni la Direction générale des entreprises (DGE), ni le Commissariat général à l'égalité des territoires, ni TECH IN France. Les parties intéressées sont invitées, le cas échéant, à faire part de leurs commentaires à la DGE.

MEMBRES DU COMITÉ DE PILOTAGE

Cédric NOZET, Direction générale des entreprises, ministère de l'Économie et des Finances
Alice MÉTAYER-MATHIEU, Direction générale des entreprises, ministère de l'Économie et des Finances
Soufiane DAHBI, Direction générale des entreprises, ministère de l'Économie et des Finances
Anne FAURE, Commissariat général à l'Égalité de Territoires, ministère de la Cohésion des territoires
Éric MENARD, responsable Études & Stratégie, TECH IN France

La conduite des entretiens et la rédaction du présent rapport ont été réalisées par la société de conseil :

ATAWAO CONSULTING

Tour Montparnasse
33, avenue du Maine
BP 05 - 75755 PARIS Cedex 15
Tél. : +33 (0)1 83 62 67 60
www.atawao.com

Consultants :

Antoine DUBOIS, directeur de mission
Demba DIALLO, directeur
Stéphan CLEMENCON, responsable du groupe S2A, maître de la chaire de ML de Telecom ParisTech
Iyed KHAMMASSI, consultant
Adama DIALLO, consultante

SOMMAIRE

Résumé	10
État de l'art technologique	13
Analyse sectorielle	16
IA, positionnement de la France et stratégies territoriales	21
Opportunités pour la France et Recommandations	25
Opportunités pour la France	25
Principales recommandations	30

RÉSUMÉ

Émergence de l'IA et contraintes

Depuis son point de départ dans les années 1950, notamment avec les travaux du mathématicien Alan TURING, l'intelligence artificielle (IA), discipline mathématique et technique destinée à reproduire l'intelligence humaine, s'est développée par cycles successifs parallèlement à la croissance de la puissance de calcul informatique disponible.

Dans les années 1980, le concept d'apprentissage automatique (« **machine learning** ») se développe permettant à une machine de déduire une « **règle à suivre** » uniquement à partir de l'analyse de données. Cette période voit apparaître la majorité des algorithmes « apprenants » utilisés aujourd'hui (réseau de neurones, apprentissage par renforcement, machines à vecteurs de support, etc.). Ces avancées se concrétisent notamment par le succès de l'ordinateur Deep Blue face au grand maître d'échecs, Gary KASPAROV en 1992.

À partir des années 2000, un nouveau cycle d'innovations en IA se met en place avec le formidable développement d'Internet et des très grandes infrastructures de calcul, offrant un accès à un volume de données encore jamais atteint dans l'histoire humaine. Avec cette capacité nouvelle, le développement des techniques d'apprentissage profond (« **deep learning** ») permet aux machines de commencer à surpasser les performances des meilleurs experts humains dans des domaines comme la reconnaissance visuelle, l'analyse documentaire ou la traduction.

Grâce à ces premiers vrais succès cognitifs des machines, une formidable dynamique mondiale d'innovation s'est mise en place depuis cinq ans avec un investissement multiplié par dix pour atteindre plus de 5 Md d'euros en 2017¹. L'innovation s'attaque à reproduire les trois grandes catégories de tâches cognitives : la perception de l'environnement, la compréhension d'une situation et la prise de décision. Le développement des technologies de vision par ordinateur ou de traitement automatique du langage (TAL) a déjà permis d'atteindre un niveau remarquable de maturité en perception et quelques succès intéressants pour des tâches de compréhension comme la traduction automatique.

Pour autant, malgré ces premiers succès, l'immense majorité des tâches de compréhension et de décision réalisées par les humains restent hors de portée des systèmes actuels en intelligence artificielle. Aucun véhicule autonome n'est aujourd'hui capable d'anticiper une situation de conduite « non apprise », alors qu'il s'agit d'une tâche relativement banale pour un humain.

La compréhension des différences entre les mécanismes d'apprentissage des machines et ceux des humains est un sujet majeur de recherche en IA. Actuellement, les performances de l'IA reposent sur la disponibilité d'un grand nombre d'événements (souvent plusieurs milliers) et une puissance de calcul importante pour l'apprentissage avec un résultat peu généralisable à d'autres situations. De son côté, le cerveau humain apprend souvent une situation avec très peu d'exemples (moins d'une dizaine) et avec une excellente capacité de généralisation à d'autres situations. Cette capacité de « généralisation » à partir de peu « d'événements » est au cœur de la recherche actuelle. Elle va de pair avec des recherches sur la « robustesse » d'un modèle apprenant, sa capacité à fournir des réponses stables malgré des événements parasites. Cela correspond à la notion « d'expérience » chez les humains.

Ces contraintes limitent l'usage de technologies d'apprentissage sans disponibilité de réseaux de communication et d'infrastructure de calcul performants. Les acteurs majeurs dans la recherche en intelligence artificielle comme GOOGLE ou AMAZON disposent des plus grandes infrastructures de calcul de la planète. Pour améliorer les performances de l'IA, la R & D matérielle (processeurs, stockage, réseau) est également très dynamique : processeur neuromorphique, réseau mobile à très faible niveau de latence, etc.

Constituer et donner accès à de très grands jeux de données de qualité

En intelligence artificielle, les technologies algorithmiques sont très largement « open source » et ne constituent pas une barrière à l'entrée. Pour l'ensemble des secteurs économiques, l'accès à des données métier de qualité (images, texte ou données numériques) est le principal facteur d'accélération ou de limitation des innovations. Une législation contraignante et une crispation globale des acteurs français (tous secteurs confondus) possédant des données, conduisent beaucoup d'innovateurs à rechercher des données d'expérimentation voire même à

¹ CB Insights.

partir à l'étranger pour développer leurs produits ou services. Sans changement ni volonté de distinguer, par exemple le besoin de données à des fins d'expérimentation/innovation, de l'usage qui pourrait en être fait après la commercialisation d'un produit ou d'un service, ce phénomène ne fera que s'accroître.

L'accès à des données massives, corrélées, complètes, qualifiées, historisées, est un **clé technologique majeure** de mise au point de technologies d'intelligence artificielle aujourd'hui. Cette donnée doit pouvoir être créée si elle n'existe pas ou rendue accessible très rapidement si elle est disponible. Sans volonté publique ou privée majeure pour lever l'ensemble des restrictions ou complications d'accès à ces données (pendant toute la phase expérimentale), aucune technologie d'intelligence artificielle ne sera développée durablement en France.

Les données existantes dans le domaine de la santé par exemple doivent être accessibles à des fins d'expérimentation de manière beaucoup plus simple et rapide qu'aujourd'hui (c'est-à-dire en quelques jours). Pour les données manquantes, l'État doit encourager la création de bases de références sectorielles (transport, industrie, énergie par exemple) pour les mettre à disposition des acteurs innovants et pour pouvoir comparer la performance des algorithmes d'intelligence artificielle entre eux. La base de données ImageNet sert de référence mondiale pour la mesure de l'évolution régulière de performance des algorithmes de reconnaissance d'image.

Capitaliser sur les usages sectoriels de l'IA

Même si la grande majorité des investissements est réalisée sur des technologies d'IA génériques, une dynamique d'usage sectoriel commence également à se développer, notamment en santé et en transport. En santé, à partir de données cliniques, des algorithmes apprenants commencent à diagnostiquer des pathologies ou imaginer un traitement personnalisé approprié. En transport, un véhicule apprend à « percevoir » et « comprendre » son environnement immédiat (signalisation, véhicules, obstacles, piétons, etc.) petit à petit pour l'ensemble des situations de conduite qui peuvent se produire : respect de la signalisation, conduite sur voie rapide, entrée sur voie rapide, embouteillage, rond-point, véhicule en double file, etc. En mobilité, l'IA commence également à apprendre à optimiser les flux de transport en fonction des capacités des différents modes, d'informations environnementales ou d'actualités (accident, grève).

Dans le secteur de l'industrie, traditionnellement utilisateur de systèmes d'optimisation, des systèmes apprenants reproduisent plus fidèlement le fonctionnement d'une machine en conditions réelles et commencent à identifier des facteurs d'utilisation non optimale ou de dysfonctionnement qui échappent aux experts humains et permettent de mieux anticiper une maintenance. Dans le domaine de la transition énergétique, l'aléa considérable des facteurs de production intermittente et de la consommation électrique constitue un terrain de jeu important pour les technologies apprenantes, notamment pour « anticiper » la production solaire ou éolienne ou encore optimiser la distribution.

Dans le secteur des affaires juridiques, des systèmes apprenants sont capables d'analyser une jurisprudence pour identifier les arguments clés, fournir un conseil juridique sur la qualité d'un contrat ou la probabilité de gagner un contentieux avec un niveau de performance comparable à un juriste professionnel. Dans le domaine de la sécurité, la reconnaissance faciale déjà performante devrait réaliser un nouveau bond en avant avec l'annonce de la Chine de développer cette technologie à partir de plus de 20 millions de caméras installées sur son territoire.

Plus globalement dans les usages numériques grand public, les grands acteurs technologiques américains ont largement investi le champ de l'IA pour cibler utilisateurs et contenus avec un niveau de précision très souvent « supérieur » aux attentes des consommateurs eux-mêmes dans le domaine. Une société comme AMAZON, retient le départ d'un camion de livraison de quelques minutes, car elle « anticipe » qu'un autre consommateur va probablement commander un produit dans les prochaines minutes dans la même zone de livraison que le camion.

Concevoir et diffuser largement en ligne des formations de référence en français

La mise au point d'algorithmes apprenants repose sur des données qui sont le reflet de phénomènes réels qui comportent une part d'aléa. Météo France prévient sur son site² que la prévision à 24 h comporte un aléa d'environ 10 %. La mise au point d'algorithmes de *Machine Learning* fait un usage important des **probabilités** dans le raisonnement. En dehors de formations spécialisées (comme l'ENSAE), cette discipline est insuffisamment enseignée de manière générale. Cela constitue une étape d'adaptation difficile à passer aujourd'hui pour toutes les équipes opérationnelles voulant travailler sur l'intelligence artificielle et développer des compétences dans le domaine.

² <http://www.meteofrance.fr/prevoir-le-temps/la-prevision-du-temps/les-performances-des-previsions>

Pour favoriser cette adaptation et développer des expertises, un effort majeur de formation doit être entrepris comme cela a été abondamment souligné dans les rapports « Stratégie France IA » et « Donner un sens à l'intelligence artificielle ». Pour accélérer cette formation, un effort particulier devrait être fait rapidement au travers de MOOCs et de projets de formations universitaires pour créer le corpus de référence en langue française sur l'IA.

Une nécessité d'être ambitieux sur l'expérimentation

Pour relever le défi de l'IA en France, le principal enjeu est de favoriser au maximum **l'expérimentation** de ces technologies sur le territoire. La notion d'expérimentation en environnement et conditions réels est clé pour fédérer largement tous les acteurs : acteurs publics, laboratoires de recherche, enseignement, écosystèmes innovants, *start-up*, grands groupes, financeurs et grand public. Ce besoin d'expérimenter est au cœur de l'initiative SIDEWALKS³ de GOOGLE par exemple avec la ville de TORONTO.

Dans le domaine de la santé, l'accès simplifié à une ou plusieurs structures de soins pour expérimenter à grande échelle des solutions innovantes accélérerait leur mise au point. Pour une structure (un laboratoire de recherche, une *start-up* ou un grand groupe) qui souhaite expérimenter, un accès simplifié signifierait : la signature d'une convention simple d'expérimentation, le déploiement rapide pendant le temps nécessaire à l'expérimentation au sein d'un service hospitalier, l'accès régulier aux experts et l'accès à des données de soins en volume.

Dans le domaine des transports et de la mobilité, une ville entière de plusieurs milliers d'habitants pourrait être transformée pour servir de champ d'expérimentation à l'ensemble des technologies innovantes du véhicule autonome. Une infrastructure pourrait être mise en place pour fournir des données et une cartographie de très grande précision, des éléments nécessaires à la mise au point de toutes les situations de conduite. En dehors d'une ville, d'autres zones d'expérimentation pourraient être mises en place comme une portion d'autoroute de plus de 100 km, une zone montagneuse, etc.

Dans le domaine de l'énergie, un ou plusieurs écoquartiers regroupant l'ensemble des technologies de transition énergétique (production énergétique intermittente, gestion intelligente de la demande, véhicules électriques, etc.) pourraient être mis à disposition des sociétés souhaitant développer des technologies d'IA dans le domaine.

Globalement, l'acceptation de l'intelligence artificielle comme facteur majeur de progrès est un enjeu de société déterminant.

³ www.sidewalklabs.com

ÉTAT DE L'ART TECHNOLOGIQUE

L'intelligence artificielle (IA) correspond à un ensemble de technologies qui permet de simuler l'intelligence et d'accomplir automatiquement des tâches de perception, compréhension et prise de décision. Ces techniques font particulièrement appel à l'utilisation de l'informatique, de l'électronique, des mathématiques (notamment statistiques), des neurosciences et des sciences cognitives.



Figure 1 - Exemple de scène à analyser⁴

Complexité croissante de l'IA	Tâche cognitive	Exemple
Mature	Perception	<ul style="list-style-type: none"> • Reconnaître le chien sur l'image • Reconnaître la route sur l'image
	Compréhension	<ul style="list-style-type: none"> • Comprendre que le chien est à côté d'une route • Comprendre que le chien est imprévisible • Comprendre que le chien pourrait sauter sur la route • Comprendre que l'on est dans une situation à risque
Recherche, exploratoire	Décision	<ul style="list-style-type: none"> • Décider de ralentir

Figure 2 - Intelligence artificielle appliquée aux tâches cognitives⁵

⁴ PXHERE.

⁵ Traitement ATAWAO.

Même si l'intelligence artificielle est principalement associée à une discipline mathématique et des techniques algorithmiques, elle inclut également d'autres briques pour répondre à un usage complet. Les principaux éléments constitutifs d'un système d'IA sont les suivants :

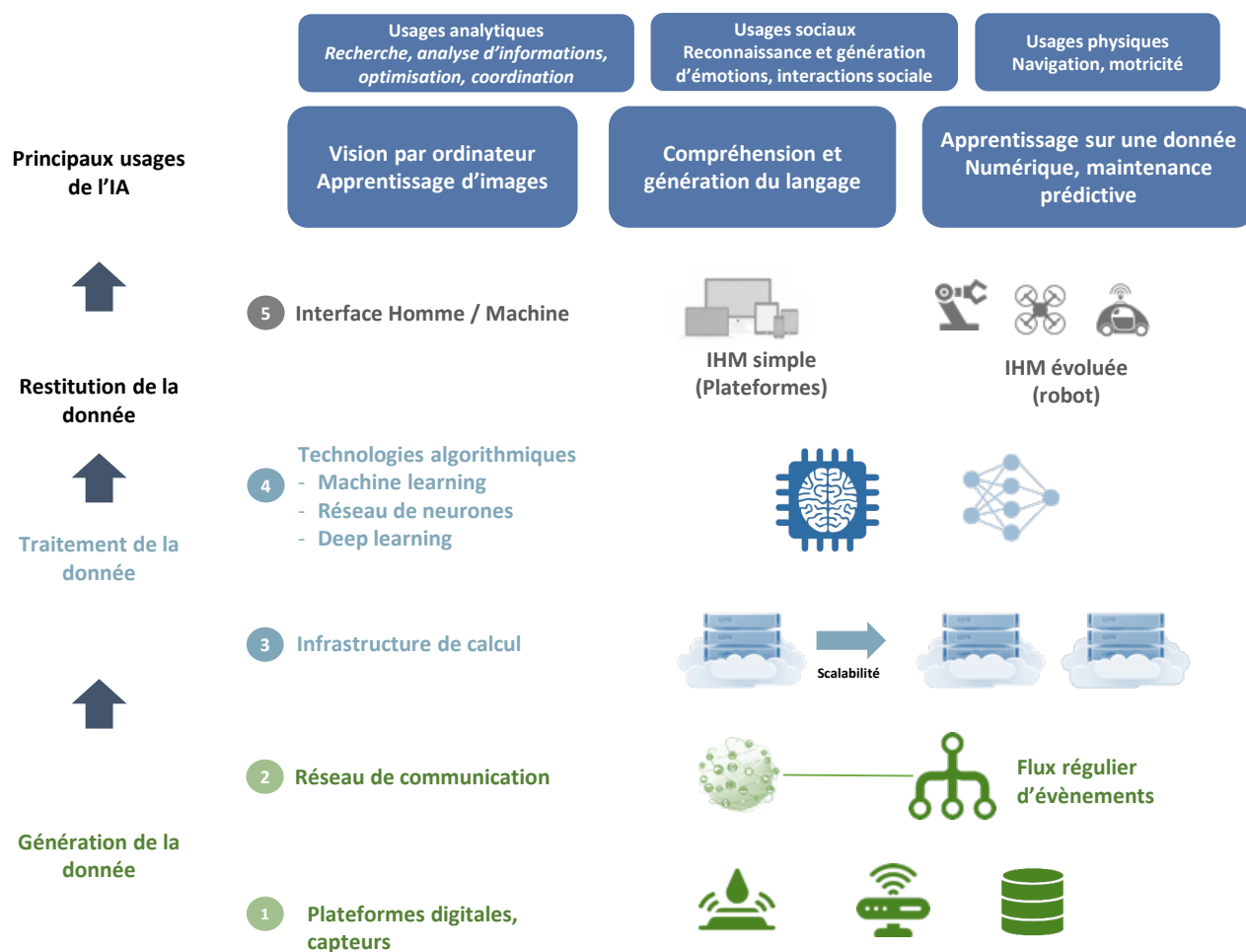


Figure 3 - Briques technologiques de l'IA

Les infrastructures de calcul fournissent les composants matériels (processeur, stockage) nécessaires aux traitements d'apprentissage et d'inférence. Les performances du stockage progressent peu depuis la mise au point des disques à état solide. À l'inverse, les processeurs font l'objet d'une dynamique d'innovation très forte pour proposer une alternative aux composants graphiques GPU massivement utilisés aujourd'hui : processeur reprogrammable en fonction des usages (FPGA), processeur « mimant » le fonctionnement d'un neurone humain (neuromorphique) ou processeur massivement multicœur à faible consommation d'énergie (KALRAY).

L'infrastructure d'échange et d'intégration de données (réseau) se déploie pour permettre la collecte croissante des données/événements utilisés par l'IA. En France, la 4G et la fibre sont en cours de déploiement. La génération suivante (5G), en cours de définition, devrait commencer à se déployer dans trois à cinq ans. L'enjeu de la prochaine génération est double : augmenter significativement les débits proposés et fournir un niveau de latence inférieur à 5 ms pour permettre des usages proches du « temps réel ». Dans cette compétition, deux groupes d'acteurs sont en concurrence : les partisans de réseaux généralistes (opérateurs) et les acteurs sectoriels qui militent pour un réseau dédié (transport, notamment).

Dans les techniques algorithmiques, les systèmes apprenants se décomposent en plusieurs couches techniques :

- Une bibliothèque « mathématique » permettant de choisir une fonction statistique, entraîner un réseau de neurones, tester sa performance et l'utiliser (SCIKIT LEARN, TENSORFLOW, CAFFE, etc.).
- Dans certains cas, un réseau de neurones déjà entraîné avec plusieurs millions de données d'un certain type (image ou texte) permettant d'atteindre des performances de reconnaissance proches des humains et utilisable en mode service (API GOOGLE vision, par exemple).

- Une bibliothèque orientée « usage » pour reproduire un tâche cognitive comme la vision par ordinateur (GOOGLE cloud vision, MICROSOFT vision API, CLARIF.ai, etc.), la compréhension du langage écrit ou parlé ou encore la traduction.

Les tâches confiées à ces algorithmes sont des tâches statistiques comme la classification, la recherche de tendances, l'identification de signal faible (exception, évènement anormal) ou la recommandation. Les usages finaux actuels découlent directement de ces tâches statistiques : *trouver un objet sur une image, reconnaître un spam ou un virus dans un flux de données, recommander un achat en fonction de ses goûts, traduire un texte, analyser un sentiment, etc.*

Aujourd'hui, la performance des algorithmes dépend presque uniquement du volume et de la qualité des données utilisées. Elles doivent posséder plusieurs caractéristiques qui dans la réalité sont très peu disponibles. Les données doivent posséder une information **représentative** du problème à résoudre. Les données de l'assurance maladie (SNIIRAM) sont représentatives de la consommation de médicaments des Français mais pas directement de leurs maladies. Elles sont peu utiles pour apprendre à faire du diagnostic automatique. Les données doivent être **labellisées** très précisément pour apprendre au système l'information cherchée. En reconnaissance d'image par exemple, tous les acteurs, GOOGLE y compris, réalisent un travail manuel colossal pour étiqueter chaque information pertinente sur une image.

La compréhension des différences entre les mécanismes d'apprentissage des machines et ceux des humains est un sujet majeur de recherche en IA. Les performances actuelles de l'IA reposent majoritairement sur la disponibilité de plusieurs milliers d'évènements et une puissance de calcul considérable pour comprendre de manière performante une situation, avec un résultat peu généralisable à d'autres situations. De son côté, le cerveau humain apprend souvent une situation avec très peu d'exemples et avec une excellente **capacité de généralisation** à d'autres situations. Cette capacité de généralisation à partir de peu d'évènements est au cœur de la recherche actuelle. Elle va de pair avec des recherches sur la **robustesse** d'un modèle apprenant et sa capacité à fournir des réponses stables malgré des évènements parasites. Cela correspond à la notion d'expérience chez les humains. **L'explicabilité** de l'IA est également un thème de recherche important. Les algorithmes d'IA sont des boîtes noires, offrant peu ou pas d'informations sur la façon dont les performances sont atteintes. L'IA explicable est un mouvement visant à développer des modèles capables de décrire leur raisonnement pour fournir un résultat.

La majorité des bibliothèques de *Machine Learning* correspond à des projets *open source* avec des enjeux scientifiques et techniques.

Pour les sociétés, l'enjeu économique porte aujourd'hui sur le développement de solutions adressant un usage transverse comme l'amélioration de la précision d'analyse des images ou du langage et de plus en plus un usage sectoriel comme la détection de fraude dans les transactions financières. Les briques technologiques de base ont été mises au point et ont donné des résultats expérimentaux intéressants. On cherche maintenant à trouver à quels problèmes l'IA peut répondre concrètement.

Un des enjeux économiques et technologiques majeurs auquel la France doit faire face réside dans le développement des infrastructures pour entraîner de grands réseaux de neurones avec des données massives. Cela nécessite des infrastructures numériques pour collecter des données à grande échelle (réseaux) et les traiter (centres de données). L'État doit encourager les opérateurs de télécommunications et les grands opérateurs CLOUD à développer leurs infrastructures.

ANALYSE SECTORIELLE

À partir d'une large revue de documents internationaux et nationaux de référence et d'une analyse approfondie de ceux-ci sur la base d'une variété de critères⁶, les secteurs considérés comme les plus impactés par l'intelligence artificielle en France aujourd'hui sont les suivants :

Rang	Résultats classement France
1	Santé
2	Industries manufacturières dont automobile
3	Transports et mobilité
4	Services d'utilité publique
5	Environnement
6	Administration publique (hors Défense)
7	Services financiers
8	Agriculture
9	Secteur juridique
10	Sécurité des biens et des personnes
11	Commerce de détail et distribution
12	Professions libérales, services professionnels
13	Éducation et recherche
14	Télécommunications et technologies
15	Loisirs et média

Tableau 1 – classement sectoriel français

La dynamique en intelligence artificielle est très différente selon les secteurs économiques. Elle est très corrélée à la digitalisation d'un secteur et donc à la disponibilité des données pour permettre l'apprentissage.

Le secteur de la santé est aujourd'hui l'un des secteurs les plus dynamiques en intelligence artificielle que ce soit en médecine préventive ou en diagnostic médical. Cette dynamique s'explique par un accès à des volumes importants de données. La **médecine préventive** bénéficie de données transmises par un nombre croissant d'objets connectés (mouvement, rythme cardiaque ou poids) et par l'effort de R & D majeur des GAFAs (APPLE et GOOGLE en tête) dans ce domaine. L'IA permet de fournir des conseils personnalisés de qualité de vie ou d'anticiper un épisode chronique (BE.CARE). **L'aide au diagnostic et au soin** est le second domaine de recherche majeur en IA notamment grâce aux données déjà majoritairement digitalisées (imagerie, électrocardiogramme ou publications médicales). L'apprentissage sur des données d'imagerie ou des électrocardiogrammes fournit déjà pour certaines pathologies comme l'arythmie cardiaque ou la mammographie, un niveau de performance supérieur à 98 % soit celui des meilleurs spécialistes humains. L'apprentissage sur des publications médicales fournit des résultats plus aléatoires pour le moment en raison, par exemple, des biais de généralisation des études cliniques ou des règles de médecine qui s'appliquent différemment d'un pays à un autre. Dans l'aide aux soins, l'IA est utilisée pour réduire les risques en contrôlant automatiquement, *via* un agent conversationnel l'état d'un patient à la place d'une infirmière ou, *via* une analyse d'image les gestes d'un chirurgien pendant une opération. Dans le domaine du contrôle des risques, les résultats sont particulièrement prometteurs. L'IA est également utilisée **en recherche clinique** pour des études épidémiologiques ou pour accélérer la mise au point de nouvelles molécules.

Pour développer les usages de l'IA en santé, l'accès à des données en volume et en qualité est une clé de performance majeure. En France, la santé est le secteur où les données disponibles sont à la fois très nombreuses par rapport à d'autres pays et en même temps le plus difficilement accessibles pour les innovateurs du fait de la législation en vigueur. C'est le frein majeur en IA aujourd'hui qui conduit les personnes souhaitant innover à se tourner vers l'étranger.

⁶ La méthodologie ayant permis d'établir ce classement repose sur deux étapes. La première consiste à établir un classement international des secteurs les plus impactés par l'IA sur la base de 18 études (70 analysées au total) et en choisissant les critères de l'innovation et de l'adoption. Une fois ce premier classement établi, la deuxième étape a consisté à adapter ce classement aux spécificités françaises en pondérant le classement par 5 critères.

Le secteur des transports et de la logistique devrait pouvoir relever des défis majeurs grâce à l'intelligence artificielle : conduite autonome, innovation de rupture dans l'optimisation de la mobilité et de la logistique. Le développement du véhicule autonome fait l'objet d'une compétition mondiale entre les constructeurs automobiles traditionnels et les géants technologiques (GOOGLE, TESLA et APPLE en tête). Plus de 80 Md de dollars ont déjà été investis depuis quatre ans, soit plus de 10 % des dépenses mondiales annuelles de R & D. La France est particulièrement en pointe dans le domaine des véhicules de transport en commun (NAVYA, EASYMILE) avec des offres déjà commercialisées. Les grands industriels du transport comme RENAULT ont également une dynamique d'innovation importante.

Dans le domaine de la logistique, les bénéfices de l'intelligence artificielle seront beaucoup plus difficiles à capter à court terme. Une chaîne logistique complète entre deux points est souvent le fait de plusieurs acteurs avec des ruptures de système d'information et donc d'accès aux données tout au long de la chaîne. De ce fait, il est déjà pratiquement impossible d'optimiser une chaîne logistique multiacteur et donc d'utiliser une intelligence artificielle pour le faire. Le partage de données entre acteurs logistiques est une condition préalable au développement de l'intelligence artificielle dans ce secteur.

Concernant le transport multimodal, plusieurs applications basées sur l'utilisation d'un calculateur prédictif pour les usagers existent. A titre d'exemple, des applications telles que CITYMAPPER, OPTYMOD'LYON ou MOOVIT permettent déjà le calcul en temps réel du meilleur itinéraire, selon différents modes de transport. WAZE l'application de navigation automobile a lancé en France, Connected Citizens, un programme d'échange de données permettant d'avoir une vision exhaustive du trafic. WAZE fournit à plus de trente partenaires (villes, entreprises) des données sur les accidents et les ralentissements signalés par les Wazers. En échange, ces derniers communiquent sur les routes fermées, les accidents, les travaux de voiries



Figure 4 - Bus EZ10 de la société EASYMILE

Pour le secteur industriel, bien que l'évolution des modèles économiques induise une remontée de la chaîne de valeur du produit vers les services, l'activité manufacturière repose toujours sur l'existence d'usines avec des procédés de fabrication. Même si l'IOT est en plein développement, la digitalisation de ce secteur s'effectue lentement. Le potentiel de gisement d'applications IA est très variable selon le type d'industrie, à la fois pour des aspects culturels, mais aussi en raison des fortes contraintes de fiabilité, sécurité et sûreté. Ces contraintes se traduisent par des réglementations strictes limitant les approches probabilistes de type « boîte noire » sur des procédés critiques.

Pour autant, la donnée machine est peut-être générée rapidement en volume important ce qui fait de l'industrie un secteur très prometteur pour l'intelligence artificielle. Des applications IA (exemples : contrôle qualité, maintenance prédictive) existent aujourd'hui déjà et sont appelées à se développer. La mutation actuelle vers l'industrie 4.0 crée de nouveaux défis à relever : nouvelles architectures industrielles, nouveaux systèmes de monitoring, nouvelles approches de régulation industrielle, intégration plus forte avec l'amont et l'aval, nouvelles exigences environnementales ou de cybersécurité.

Les secteurs de l'énergie et de l'environnement constituent dès à présent des lieux d'application à fort impact de l'IA. C'est le cas, par exemple, des apports de l'IA pour : le pilotage des réseaux électriques intelligents, les stratégies d'efficacité énergétique ou encore l'anticipation des chocs météorologiques ou de celui du changement climatique.

La France recèle de nombreux atouts dans ces domaines : outre la présence sur le sol national de grands groupes *leaders* mondiaux (EDF, ENGIE, LEGRAND, VEOLIA...), ainsi que de PME (exemple : SOCOMEC) et de *start-up* innovantes (exemples : DC BRAIN, ENERGIENCY) sur toute la chaîne de valeur, le pays compte aussi de grands centres de recherche publics (CEA, par exemple). Atouts qui se sont traduits pour l'heure dans la mise en œuvre de nombreux projets et démonstrateurs intégrant des composantes d'IA. Cette base solide pourrait constituer le socle d'une ambition nationale plus forte et plus ciblée de l'IA pour l'énergie et l'environnement, dans le cadre de stratégies partenariales public-privé à amplifier par la dimension européenne.

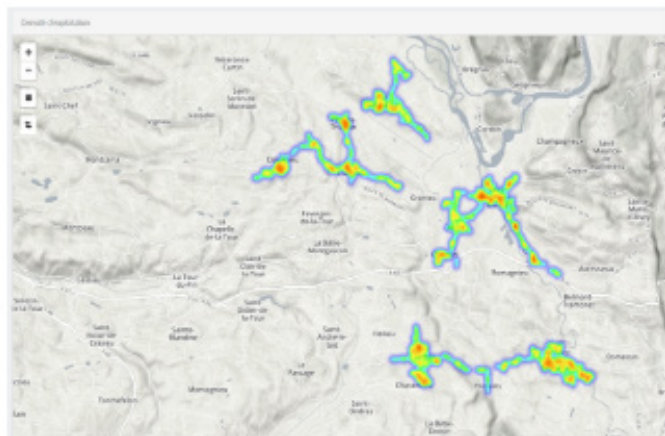


Figure 5 - DC Brain, fuites et anomalies réseau

En dehors de ces quatre secteurs, les plus matures sont aujourd'hui **le commerce, les services financiers et la sécurité des biens et des personnes**. Ces secteurs se caractérisent par un accès à de grandes quantités de données, des technologies déjà matures pour adresser des usages commerciaux et des sociétés commercialisant déjà des produits dans plusieurs pays.

Le secteur du **commerce en ligne** est le principal secteur d'application de technologies d'apprentissage au niveau mondial. Il concentre un accès à des données illimitées, les investissements majeurs de géants du Web comme AMAZON et ALIBABA et un objectif très clair de développement des ventes. AMAZON a créé un institut de recherche de plus de 100 personnes en intelligence artificielle à TURBIGEN à côté de l'institut Max Planck. La majorité des usages en commerce correspond à un « matching » entre un produit et un client potentiel. À l'inverse, si on se limite à la France, le secteur est très peu mature en intelligence artificielle. Aucun acteur de commerce en ligne ou non français n'est à même de rivaliser avec les Léviathans mondiaux.

Le secteur des **services financiers**, très digitalisé, est également l'un des plus matures dans les technologies de *Machine Learning*. Les usages se déploient en *trading* (analyse prédictive ou *trading* automatisé), pour les opérations courantes (*scoring* client, conformité, fraude) ou les services de support à la clientèle. Ces services sont majoritairement fournis par des acteurs étrangers. Le secteur de la **sécurité des biens et des personnes** est également très dynamique car beaucoup de services reposent sur de l'analyse d'images, l'un des domaines les plus matures pour le *machine learning* et le *deep learning*. L'analyse automatique de scène de vidéosurveillance est le thème de recherche qui se développe le plus vite. La France possède l'un des *leaders* mondiaux dans ce domaine, la société IDEMIA (ex-MORPHO).



Figure 6 - Reconnaissance faciale pendant un entretien vidéo⁷

En dehors des secteurs matures, plusieurs secteurs émergent également. Ces secteurs sont dans une phase expérimentale des usages de l'IA à partir des technologies mises au point. Les sociétés qui innovent adressent déjà un premier cas d'usage, mais cherchent encore leur modèle économique. Il s'agit des secteurs des professions juridiques, des loisirs, des télécommunications, des services professionnels et des administrations publiques. Pour **les professions juridiques**, les usages se développent en droit commercial (anglo-saxon) pour préparer, analyser et évaluer automatiquement les risques d'un contrat. La jurisprudence est largement digitalisée, ce qui permet de regrouper rapidement les éléments juridiques constitutifs d'un cas : création d'un contrat, analyse des clauses ou préparation de contentieux. Les sociétés du secteur juridique et technologique (LegalTech) ne sont pas encore très développées en France. Dans le secteur des **loisirs**, ce sont les mêmes usages que le commerce qui se développent. Majoritairement, ils permettent d'associer un contenu multimédia à un consommateur pour proposer une expérience interactive très personnalisée. Dans le domaine des télécommunications, les usages se développent sur l'optimisation des ressources réseau ou la maintenance prédictive des équipements. Dans les **services professionnels**, trois types d'usages se développent : l'évaluation ou la détection d'un risque, l'assistant de recherche ou de formation personnalisé ou l'optimisation d'un mode de travail. L'enjeu d'efficacité est toujours important dans les entreprises, ce qui devrait favoriser les usages de l'IA. Les usages se développent plus lentement car les données sont plus fragmentaires et la question d'optimisation est plus compliquée à adresser. Dans les **administrations publiques**, les usages se développent pour évaluer plus rapidement l'impact d'une politique publique ou pour répondre aux questions des citoyens.

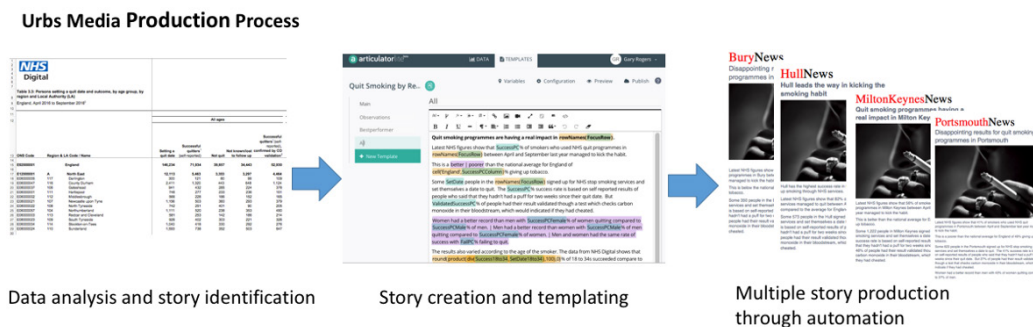


Figure 7 - Processus de production d'URBS MEDIA

L'application de l'intelligence artificielle est plus difficile à évaluer pour **l'agriculture**. En dehors de la météorologie, le manque de très grands jeux de données annotées aisément accessibles rend difficile pour le moment l'application de techniques d'apprentissage, même s'il existe des enjeux importants de conservation des ressources naturelles. Par ailleurs, la superficie importante rend également le travail de collecte de données, *via* un réseau de communication, plus difficile.

⁷ Hirevue.

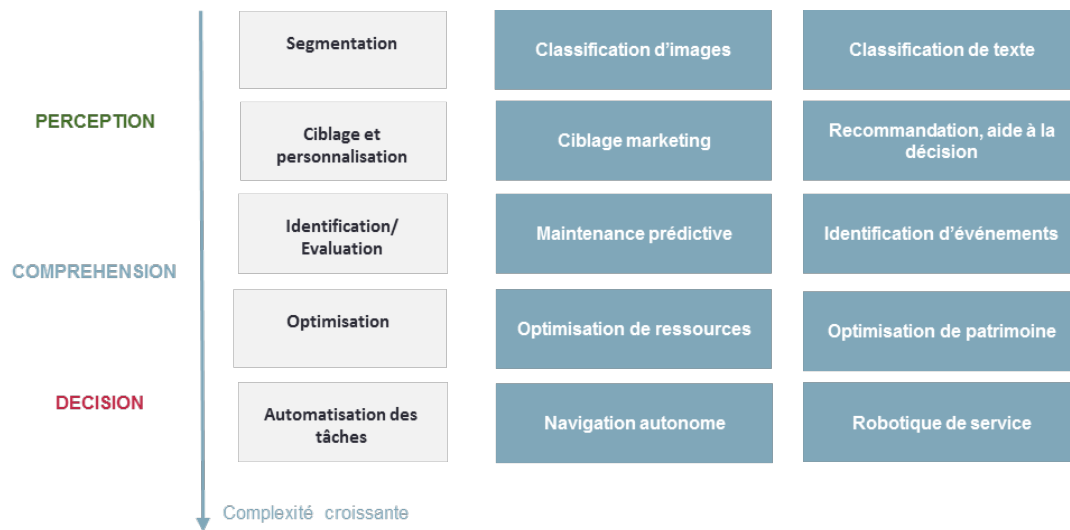


Figure 8 - Typologie générale des usages IA

IA, POSITIONNEMENT DE LA FRANCE ET STRATEGIES TERRITORIALES

L'IA est un sujet émergent. Analyser ses impacts sur les territoires, revient à analyser dans les années 90, l'impact spatial de la diffusion de l'internet mobile. Les acteurs interrogés n'ont pas encore de vision des impacts sur les territoires de la diffusion de l'IA ou sur l'attractivité d'entreprises en IA. Vouloir analyser de manière très spatialisée les effets de l'IA nécessiterait des études précises et localisées.

En conséquence, pour étudier le lien entre IA et territoires, l'approche retenue permet d'illustrer quelques expériences territoriales emblématiques à l'échelle mondiale et de faire le parallèle avec le territoire national. La méthode utilisée est la suivante :

- Etude des territoires les plus compétitifs à l'échelle mondiale sur les sociétés innovantes et les talents en IA.
- Illustration de l'attractivité de talents et d'investissements de ces territoires.
- Identification des usages innovants d'IA pour comprendre quelles actions locales peuvent être entreprises.
- Etablissement des principaux constats et des recommandations associées.

3 645 *start-up* appartiennent au domaine de l'IA dans le monde dont 40 % aux États-Unis, 11 % en Chine, 10 % en Israël et 7 % au Royaume-Uni. Avec 109 *start-up*, la France est au septième rang mondial (3,1 %) proche du Canada, du Japon et de l'Allemagne⁸.

Une ambition atteignable serait d'arriver au niveau du Canada, voire de dépasser celui-ci. Ce pays avec une économie et une population moins importante que la France possède 2 villes dans le peloton de tête mondial.

Hors métropole francilienne, plusieurs grandes villes commencent à s'organiser sur la thématique IA (Lyon, Lille, Marseille, Toulouse...). Avec l'aide de la puissance publique pour amplifier les efforts et dans le cadre d'une stratégie bien ciblée, Lyon pourrait rejoindre le peloton de tête mondial.

Le fait d'apparaître dans le classement de tête de la localisation de l'innovation IA (représentée notamment par le volume de startups spécialisées sur la thématique) a pour vertu d'envoyer un signal aux différents acteurs (investissements, talents) et de contribuer à les attirer.

Une nette domination américaine

Cette suprématie américaine se retrouve au niveau de la concentration de *start-up* dans deux villes : San Francisco, premier pôle IA au monde (17 %) et New York, quatrième pôle (5 %), c'est-à-dire près de six fois et deux fois plus que la France. Les usages IA les plus importants sont développés par des entreprises situées à San Francisco et dans la Silicon Valley (GOOGLE, FACEBOOK, TWITTER, LINKEDIN, etc.) et dans la région de New York également, du fait des applications dans la publicité, le *marketing* ou les FINTECH par exemple.

Les universités produisant le plus grand nombre d'entrepreneurs dans le numérique sont aussi en Californie (Stanford, Berkeley). Les montants investis par les fonds d'investissement et les moyens dont disposent celles-ci sont sans commune mesure à l'échelle mondiale. La recherche en IA est également très active dans les agences de l'administration fédérale de l'énergie et de la défense avec la *Defense Advanced Research Projects Agency* (Darpa) qui a initié le programme *AI NEXT*⁹ de 2 Md de dollars ou la Nasa¹⁰ qui dispose d'entités dans les principaux pôles technologiques du pays dont la Silicon Valley.

Les États-Unis attirent les meilleurs talents de la planète avec l'attractivité de leurs universités et des salaires dans les entreprises IA, les plus élevés au monde.

⁸ Artificial Intelligence – A strategy for European startups, ROLAND BERGER, ASGARD, 2018

⁹ <https://www.darpa.mil/work-with-us/ai-next-campaign>

¹⁰ <https://www.nasa.gov/feature/goddard/2017/nasa-explores-artificial-intelligence-for-space-communications>

La Chine, second pôle mondial : une stratégie affirmée à l'échelle des territoires

La Chine protège son marché intérieur en limitant sur son territoire les services des GAFA, ce qui a permis un développement de près de 383 *start-up* IA dont une partie est à même d'assurer la compétition au niveau mondial comme CLOUDWALK spécialiste de la reconnaissance faciale, CAMBRICON, SENSETIME¹¹, UBTECH ROBOTICS ou DJI.

Conformément à la vision 2030 de l'État Central, les gouvernements de 18 provinces et municipalités ont publié leurs stratégies pour promouvoir les écosystèmes IA au niveau local¹². Le gouvernement de Tianjin prévoit de créer un fonds de 16 Md de dollars pour soutenir le développement de l'IA¹³.

Plusieurs pôles IA émergent. Beijing abrite 43 % des *start-up* IA chinoises. Elle bénéficie d'institutions de recherche de renom comme les universités de TSINGHUA ou BEIHANG. GOOGLE a récemment ouvert un centre d'IA à Beijing. Shenzhen accueille 20 % des *start-up* IA chinoises. Elle bénéficie de grandes entreprises comme TENCENT ou HUAWEI par exemple.

Tel-Aviv, troisième pôle mondial, la moitié des *start-up* israéliennes

Israël et ses principales villes dont Tel Aviv et ses grandes universités comme le TECHNION sont très connectés aux États-Unis. Créée en 2015, LEMONADE, *start-up* israélienne à cheval entre Tel-Aviv et New York dans l'Insurtech vaut à présent près de 500 millions de dollars. L'IA est au cœur de ses services.

L'écosystème innovant israélien est dynamisé par la présence de plus de 300 centres de R & D, de multinationales, de liens étroits entre l'armée (avec de nombreux usages dans l'analyse de données et la sécurité) et le monde académique et d'une forte culture entrepreneuriale. Israël a pu prendre une avance considérable sur le marché de l'IA. Ce pays compte 362 *start-up* IA (52 % à Tel-Aviv), soit 10,4 % du total mondial, un chiffre très proche de celui de la Chine (11 %).

Londres, plus que Paris et Berlin réunis

Londres possède le nombre le plus important de *start-up* européennes et bénéficie de l'effet d'aubaine de la présence d'un *leader* mondial DEEPMIND, entreprise rachetée par GOOGLE. La société dispose d'un effectif en IA unique au monde et contribue à créer des effets d'entraînement (développement de *start-up* dans son sillage). Londres compte plus de 200 *start-up* IA, davantage qu'à Paris (73), Berlin (56) et Helsinki (32) réunis.

La France et Paris, premier rang européen continental pour l'attractivité en IA

D'après le rapport Roland Berger-Asgard¹⁴, la France est la première nation continentale européenne et Paris la première ville européenne, en termes d'attractivité des *start-up* IA. La France arrive avant l'Allemagne (huitième) et Paris devant Berlin (douzième).

Paris localise les deux tiers des *start-up* (73 sur 109). C'est aussi à Paris que se sont localisés les centres de R & D en IA de FACEBOOK ou de GOOGLE par exemple. En mars 2018, le CNRS, L'Inria, l'Université PSL, AMAZON, CRITEO, FACEBOOK, FAURECIA, GOOGLE, MICROSOFT, NAVER LABS, NOKIA BELL LABS, PSA, SUEZ et VALEO, se sont associés pour créer l'Institut PRAIRIE pour *PaRis Artificial Intelligence Research InstitutE*¹⁵. Plus largement, 45 % des laboratoires français publics et privés, disposant de compétences IA sont localisés en Île-de-France.

Cette dynamique autour de l'IA en Île-de-France, se traduit par une augmentation des demandes de profils qualifiés. Selon l'Association pour l'emploi des cadres (Apec), le nombre d'offres d'emploi pour des projets d'intelligence artificielle a doublé en 2017¹⁶ avec 2 398 postes dont 63 % des entreprises en Île-de-France. L'annonce récente faite par la région Île-de-France du lancement du Plan « IA 2021 » devrait renforcer cette polarisation autour de Paris et de sa région¹⁷. Cet état de fait n'est pas anormal, car même à l'échelle de très grands pays comme les États-Unis ou la Chine, les principaux pôles d'attractivité de l'innovation IA sont limités (sept aux États-Unis et deux en Chine). On peut toutefois s'interroger sur un modèle « à la française » qui verrait travailler en réseau les spécialistes IA des quinze pôles industriels français d'envergure mondiale (voir <http://competitivite.gouv.fr/>).

¹¹ <https://www.sensetime.com/>

¹² <https://www.merics.org/cn/node/6851>

¹³ <https://www.reuters.com/article/us-china-ai-tianjin/chinas-city-of-tianjin-to-set-up-16-billion-artificial-intelligence-fund-idUSKCN11I0DD>

¹⁴ Artificial Intelligence – A strategy for European startups, ROLAND BERGER, ASGARD, 2018

¹⁵ http://www2.cnrs.fr/sites/communique/fichier/cp_prairie_2018_03_30vf.pdf

¹⁶ https://presse.apec.fr/files/live/mounts/media/fichiers/Note_Apec_juin%202018_Intelligence_artificielle%20v11.pdf

¹⁷ <https://www.iledefrance.fr/presse/la-region-ile-de-france-presente-plan-regional-l-intelligence-artificielle-ia-2021-les>

Les usages de l'IA qui se développent sur les territoires

Smart Cities. GOOGLE a lancé une initiative structurante sur la thématique « smart cities » en signant avec la ville de Toronto un projet de construction d'un nouveau quartier¹⁸ sur une friche portuaire de 5 hectares. Le contrat, confié à SIDEWALK LABS¹⁹ sera opérationnel à horizon 2022 et pourrait créer jusqu'à 5 500 emplois. Ce nouveau quartier écologique abritera le siège social canadien de GOOGLE et utilisera massivement la collecte des données pour des services autour de la mobilité urbaine ou la gestion intelligente des bâtiments et de l'énergie.

Prise en charge des seniors par des robots. Au Japon, l'IA et la robotique sont au cœur des réflexions de la ville de Yokohama, par exemple, pour répondre aux besoins spécifiques des seniors qui représenteront 30% de la population de la ville en 2025. Déjà de grands groupes comme FUJITSU²⁰ ont développé une solution à base de capteurs multiples et d'IA pour rendre les robots soignants plus adaptatifs et faciliter le travail du personnel soignant.

Mobilité urbaine grâce aux voitures autonomes. San Francisco a instauré en 2014 un cadre réglementaire qui permet aujourd'hui à plus de 50 sociétés (AURORA²¹, ZOOX,²² TESLA, GENERAL MOTORS, TOYOTA, VOLKSWAGEN GOOGLE, APPLE, UBER...) d'entamer des essais de conduite de voitures autonomes²³. Les besoins économiques et les fortes contraintes géographiques ont conduit Singapour à élaborer le Singapore Autonomous Vehicle Initiative (SAVI)²⁴ pour pouvoir sécuriser l'introduction des véhicules autonomes.

Transport multimodal avec un calculateur prédictif pour les usagers. Les applications telles que CITYMAPPER, OPTYMOD'LYON²⁵ ou MOOVIT permettent le calcul en temps réel du meilleur itinéraire, selon différents modes de transport. WAZE l'application de navigation automobile a lancé en France, Connected Citizens,²⁶ un programme d'échange de données permettant d'avoir une vision exhaustive du trafic. WAZE fournit à plus de trente partenaires (villes, entreprises) des données sur les accidents et les ralentissements signalés par les Wazers. En échange, ces derniers communiquent sur les routes fermées, les accidents, les travaux de voiries...

Optimisation des services hospitaliers. Les Hospices civils de Lyon ont lancé un projet avec MICROSOFT, comportant différents objectifs auxquels contribueront des technologies d'IA : aide au diagnostic, simplification du travail des soignants, médecine prédictive²⁷.

L'IA, un sujet encore peu intégré à date, aux stratégies de développement économique et d'innovation des territoires

La France ne fait pas exception aux constats que l'on peut faire à l'échelle internationale, au sein de grandes nations industrielles ou émergentes : les développements récents et les promesses de l'IA étant des opportunités assez nouvelles et peu matures, les politiques publiques dédiées, commencent seulement elles aussi, à apparaître. Elles sont conçues à l'échelle des pays (politiques publiques nationales) et existent peu aux échelles locales.

En effet, à l'échelle mondiale, à part quelques grandes régions (la région Ile de France par exemple ou les exemples chinois cités plus haut) et de très grandes métropoles (Londres comme illustré plus haut), très peu de territoires ont conçu des politiques publiques spécifiquement tournées vers la promotion de l'IA.

En France, le constat est le même : la région Ile de France est la seule, à date, à avoir conçu des politiques publiques spécifiquement dédiées à la diffusion des usages IA et à l'attraction de talents et d'entreprises sur cette thématique.

¹⁸ https://www.itespresso.fr/smart-cities-google-toronto-187066.html?inf_by=5bba5cfe671db8fb278b4e75

¹⁹ <https://www.sidewalklabs.com/>

²⁰ <http://www.fujitsu.com/sg/about/resources/news/press-releases/2017/fcal-20172103.html>

²¹ <https://aurora.tech/>

²² <https://zoox.com/>

²³ https://www.lemonde.fr/economie/article/2018/02/28/la-californie-ouvre-la-voie-aux-taxis-autonomes_5263565_3234.html

²⁴ <https://www.lta.gov.sg/content/ltaweb/en/roads-and-motoring/managing-traffic-and-congestion/intelligent-transport-systems/savi.html>

²⁵ <https://www.cityway.fr/realisation/optymodlyon-calculateur-multimodal/>

²⁶ <https://www.waze.com/fr/ccp>

²⁷ <https://www.ladepeche.fr/article/2018/05/30/2807871-les-hopitaux-de-lyon-testent-l-intelligence-artificielle.html>

Par ailleurs, l'examen des différents instruments de développement économique des territoires que sont le SR2I (Schéma régional de développement économique d'innovation et d'internationalisation), les pôles de compétitivité ou encore les TIGA (Territoires d'Innovation- Grande Ambition), n'a pas permis de dégager des orientations fortes des régions en matière d'IA. Ces différents dispositifs intègrent tous le numérique, comme levier d'innovation et de développement économique. A titre d'exemple, plusieurs pôles de compétitivité sont spécialisés autour de cette thématique et comportent probablement divers projets collaboratifs intégrant de l'IA.

Compte-tenu du potentiel de forte transformation apporté par l'IA sur un grand nombre d'activités, il est souhaitable que les territoires s'emparent spécifiquement de cette thématique et mettent en œuvre des diagnostics visant à identifier leurs forces et faiblesses avant de concevoir des stratégies adaptées.

Il sera également nécessaire de fédérer et coordonner ces initiatives locales. L'IA française n'est pas d'une taille suffisante pour permettre une dispersion des initiatives. L'IA requiert, quelles que soient les entreprises, les localisations ou les thématiques concernées, le regroupement de moyens et la réalisation d'économies d'échelles. Qu'il s'agisse de la mise à niveau des ressources de calcul intensif ou de cloud, de l'organisation de rencontres d'envergure internationale, de la structuration de la formation, et avant tout l'enrichissement du patrimoine de ressources immatérielles, le développement de l'IA française doit être organisé sur un mode combinant la recherche de l'excellence et la mise en réseau des acteurs les plus performants.

Ainsi il apparaît essentiel de favoriser l'émergence d'une démarche fédérative de l'IA française. A l'instar des exemples étrangers identifiés plus haut, il faudra associer les collectivités territoriales et notamment impliquer les Régions, qui financent une grande partie de l'effort de recherche et d'industrialisation. La dimension IA devra être prise en compte plus systématiquement dans les grandes politiques de soutien au tissu entrepreneurial et économique et ainsi être intégré par exemple dans la dynamique du programme « Territoires d'industrie ».

OPPORTUNITÉS POUR LA FRANCE ET RECOMMANDATIONS

Opportunités pour la France

L'étude sectorielle et l'analyse des différentes caractéristiques techniques associées à l'IA permettent de retranscrire, sous forme de schéma, les domaines privilégiés d'application de l'IA en France ainsi que leurs prérequis techniques.

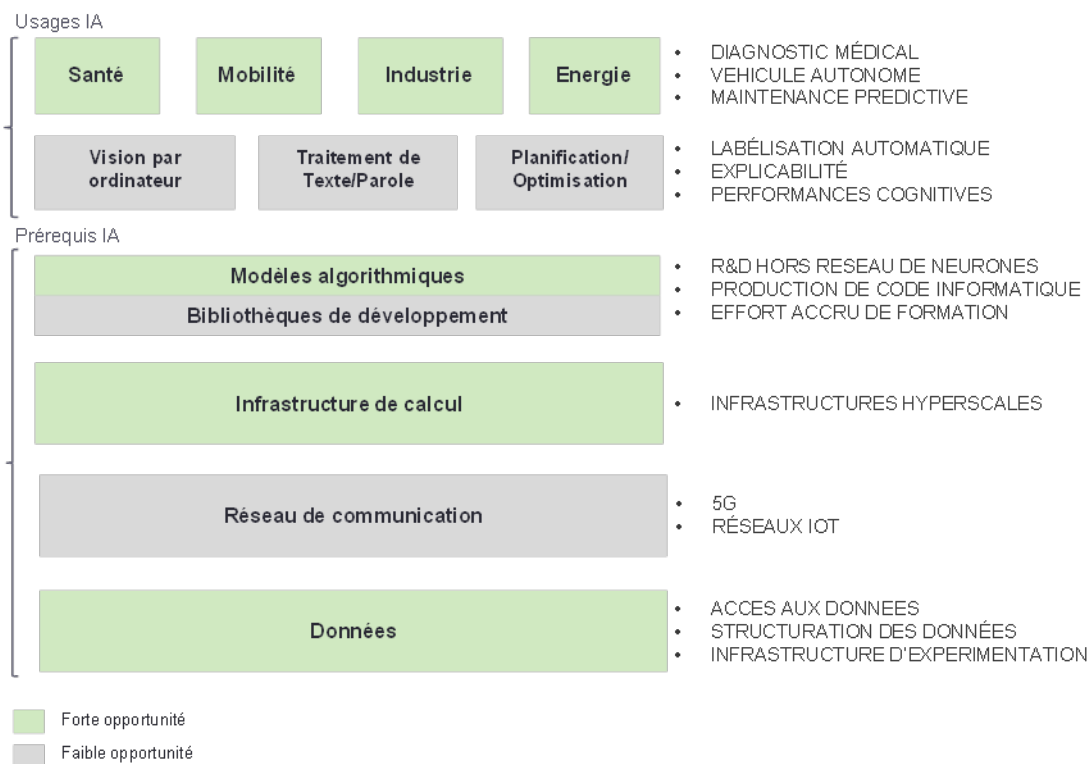


Figure 9 - Opportunités pour la France

Développer et sécuriser la recherche en intelligence artificielle en France

La révolution de l'intelligence artificielle en est encore à ses prémices. Les techniques développées jusqu'à maintenant ne soutiennent pas la comparaison avec les performances cognitives évoluées des humains. Les progrès actuels, spectaculaires pour certains usages très spécifiques, ne peuvent masquer l'immense chantier de recherche à produire sur les vingt prochaines années pour reproduire efficacement les capacités d'apprentissage uniques des humains.

C'est l'une des premières opportunités pour la France. L'enjeu est double : former et retenir en France les meilleurs chercheurs en intelligence artificielle. Cela suppose de développer des efforts dans plusieurs domaines :

- Prioriser les investissements sur des champs de recherche actifs (apprentissage avec peu de données, robustesse des modèles face à la perturbation, explicabilité, etc.).
- Stimuler les collaborations transdisciplinaires dans les appels à projets, car l'IA est une discipline qui nécessite une forte collaboration entre expertise scientifique, métier et informatique.
- S'assurer que les aides et dispositifs fiscaux de soutien à la R & D ciblent des verrous scientifiques et technologiques prioritaires et non des travaux d'ingénierie sans dépassement de l'état de l'art.

- Valoriser la recherche appliquée en stimulant par exemple la production d'algorithmes documentés en support à la recherche et pas seulement de publications scientifiques.
- Inciter les entreprises privées à compléter fortement les financements publics insuffisants en IA. Lorsque le MIT annonce une décision d'investissement de 1 milliard de dollars dans l'IA²⁸, – on comprend qu'elle ne peut pas être traitée à égalité avec les autres thématiques académiques.
- Encourager le secteur privé à mettre en place une R & D à long terme sur l'intelligence artificielle. Il dispose des données, de l'expertise métier, et des enjeux à adresser. En développant une R & D forte ce sujet, il aidera à faire monter en puissance la recherche académique.
- Mettre en lumière (visibilité dans les grands congrès scientifiques, développement de chaires de recherche, etc.) l'élite des centres de recherche français pour structurer autour d'eux des équipes de pointe dans le domaine.
- Inscire la trajectoire de la recherche française dans une perspective européenne pour mutualiser les efforts. Il s'agira ici non seulement de s'inscrire dans des lignes directrices proposées par la Commission européenne pour générer une masse critique d'efforts de recherche, mais aussi d'approfondir certains sujets avec des pays volontaires. Compte tenu de l'importance du Royaume-Uni sur la thématique IA en Europe, il est crucial que ce pays fasse partie de cette alliance.

²⁸ <https://www.technologyreview.com/the-download/612293/mit-has-just-announced-a-1-billion-plan-to-create-a-new-college-for-ai/>

Généraliser les offres de formations

Les technologies IA sont des GPT (*General Purpose Technologies*), c'est-à-dire, des technologies transverses à effet de rupture comme la machine à vapeur, l'électricité, Internet... Aujourd'hui, les technologies étant majoritairement en *open source* donc accessibles, le défi le plus urgent est dans le développement des compétences qui permettraient de développer ce patrimoine de ressources.

De ce fait, il est crucial pour accompagner leur diffusion, d'intensifier la formation du plus grand nombre d'effectif possible pour que chaque secteur dispose de compétences pertinentes pour ses propres applications.

La France dispose d'atouts importants : grand nombre et variété des écoles d'ingénieur, enseignement mathématique d'excellent niveau, autres formations pouvant développer des compétences de développement logiciel et de *Machine Learning* (ex. : formations courtes de type BTS ou IUT pour les premiers niveaux, universitaires généralistes dans des disciplines scientifiques pour des formations plus avancées...).

Se focaliser sur les applications sectorielles de l'IA

Les technologies algorithmiques de base en intelligence artificielle sont largement *open source* et accessibles à l'ensemble des innovateurs. Elles sont mises au point dans les laboratoires de recherche académiques comme l'Inria avec SciKit-Learn et contribuent au développement des connaissances. Ces technologies concernent tout ce qui est transversal : traitement du texte, de l'image, transformations mathématiques, modélisation des réseaux de neurones, etc. Encourager le développement de ces technologies répond essentiellement à des enjeux scientifiques.

Les enjeux économiques se portent de plus en plus sur des usages sectoriels de l'intelligence artificielle : diagnostic médical, maintenance préventive, détection automatique de piéton, vidéosurveillance automatisée, aide à la vente, etc.

Pour les secteurs grand public (bien-être et prévention en santé, moteur de recommandation pour le *Retail*, moteurs de recherche, assistants vocaux, etc.), compte tenu de la domination des GAFAM/BATX (ex. : AMAZON a une équipe de plus de 5 000 personnes sur les assistants ALEXA et ECHO²⁹) et de l'importance des moyens investis, la France ne dispose pas d'atouts spécifiques. Il n'est pas impossible que des acteurs nationaux de premier plan apparaissent, mais la probabilité est faible et les besoins de ressources conséquents. En conséquence, l'État doit focaliser son accompagnement et ses ressources sur des secteurs B2B ou B2B2C. À l'échelle mondiale, le niveau de maturité des acteurs est sensiblement équivalent dans les secteurs B2B.

L'IA pour la gestion de l'attrition client ou de la fraude dans le secteur bancaire ou pour la maintenance préventive dans l'industrie possède un niveau de maturité proche de celui d'un voisin européen ou d'un acteur nord-américain. Sur ces segments, la France dispose déjà de grands acteurs établis (DASSAULT SYSTEMES) ou récents (CRITEO), mais surtout d'un écosystème prometteur de *start-up* de premier plan (DATAIKU, SHIFT TECHNOLOGY, NAVYA, EASYMILE, STANLEY ROBOTICS, PROPHESEE, etc.).

À l'aide de la BPI, il faut structurer autour de ces *start-up* un écosystème complet (autres *start-up*, grandes entreprises, centres académiques, clients...) pour accroître les passerelles et les coopérations entre les acteurs, rendre visible et attractive les filières d'excellence française et mettre en lumière un potentiel pour le recrutement.

Encourager massivement le développement d'infrastructures numériques

Les acteurs majeurs dans le domaine de l'Intelligence artificielle sont aussi ceux qui disposent des plus grandes infrastructures de calcul à leur disposition et ce n'est pas une coïncidence. Dans la course à l'automatisation de tâches cognitives avec des performances comparables aux humains, les *leaders* s'appuient sur des infrastructures qui permettent un passage à l'échelle des modèles de réseaux de neurones mis au point. Ces infrastructures permettent d'entraîner les modèles avec plusieurs millions de données dans un temps acceptable.

²⁹ <https://www.cnn.com/2017/09/27/amazon-has-5000-people-on-echo-and-alexa-more-than-fitbit-and-gopro.html>

L'Europe et la France en particulier, accusent un **retard très important sur l'utilisation de grandes infrastructures de calcul**. Très peu d'acteurs ont compris l'importance stratégique de ce sujet technologique. Pour faire une analogie avec la géographie, les États-Unis, l'Europe ou la Chine représentent chacun autour de 7 % de la superficie du globe. En géographie numérique représentée par les centres de données *hyperscales*, sur 400 infrastructures³⁰ de ce type dans le monde, 44 % sont aux États-Unis, 20 % en Europe, 8 % en Chine et moins de 1 % en France. Les États-Unis numériques représentent plus de six fois leur taille géographique réelle.

Au-delà de ces chiffres, encourager massivement le développement des acteurs d'infrastructures numériques en Europe et en France est fondamental pour comprendre et maîtriser le résultat de l'entraînement des modèles. La compréhension des mécanismes d'entraînement, c'est-à-dire l'explicabilité, permet de vérifier si une intelligence artificielle respecte les contraintes, les lois, les règles, les règlements ou l'éthique. Dans le cas contraire, l'IA devient une boîte noire, sans que l'on puisse identifier quelles données ont servi à l'entraînement et si ces données respectent toutes les contraintes de diversité, complétude ou précision nécessaire. Un centre de données *hyperscales* est le lieu où sera mis au point l'ensemble des intelligences artificielles de production.

Lancer des expérimentations ambitieuses

Le principal objectif des actions proposées est de favoriser au maximum **l'expérimentation** de technologies d'intelligence artificielle sur le territoire. La notion d'expérimentation en environnement et conditions réelles est clé pour fédérer largement tous les acteurs impliqués : acteurs publics des administrations centrales et territoriales, laboratoires de recherche, enseignement, écosystèmes innovants, *start-up*, grands groupes, financeurs et grand public. Ce besoin d'expérimentation à grande échelle est notamment à l'œuvre dans le cadre de la mise au point d'un véhicule autonome et mobilise déjà largement l'ensemble des acteurs en ordre plus ou moins dispersé. C'est une dynamique qui est pertinente à reproduire dans le cadre d'autres domaines d'application de l'IA. Ces zones d'expérimentation offriraient un cadre administratif et légal, une infrastructure produisant de la donnée en volume et en flux et une organisation permettant de mettre en application l'IA (utilisateurs, experts, etc.).



Figure 9 - Robot Pepper en expérimentation dans un hôpital

Dans le domaine de la santé par exemple, l'accès simplifié à une ou plusieurs structures de soins pour expérimenter à grande échelle des solutions innovantes accélérerait leur mise au point. Pour une structure qui souhaite expérimenter (un laboratoire de recherche, une *start-up* ou un grand groupe), un accès simplifié signifierait : un processus simple de signature d'une convention d'expérimentation, le déploiement rapide pendant le temps nécessaire à l'expérimentation au sein d'un service hospitalier, l'accès régulier aux experts métiers et à des données de l'hôpital. Pour améliorer l'accès aux données, une infrastructure pourrait compléter les systèmes existants pour garantir la disponibilité d'un maximum de données.

Dans le domaine des transports et de la mobilité, une ville entière de plusieurs milliers d'habitants pourrait être transformée pour servir de champ d'expérimentation à l'ensemble des technologies innovantes du véhicule autonome. Une infrastructure pourrait être mise en place pour fournir des données et une cartographie de très grande précision, nécessaire à la mise au point de toutes les situations conduite. En dehors d'une ville, d'autres zones d'expérimentation pourraient être mises en place comme une portion d'autoroute de plus de 100 km, une zone montagneuse, etc.

³⁰ Synergy Research Group 2018.

Dans le domaine de l'énergie, un ou plusieurs écoquartiers regroupant l'ensemble des technologies de transition énergétique (production énergétique intermittente, gestion intelligente de la demande, véhicules électriques, etc.) pourraient être mis à disposition des sociétés souhaitant développer des technologies d'IA dans le domaine.

Faciliter un accès rapide aux données

Le second objectif des actions vise à **rendre le plus accessible possible toutes les données** disponibles pendant la phase de recherche et de développement de technologies d'intelligence artificielle. Les actions proposées visent à créer une distinction explicite et légale entre la phase d'expérimentation et celle de l'usage commercial de données. Cette distinction peut permettre de simplifier considérablement l'accès aux données pendant les phases de mise au point.

La facilité d'accès aux données représente un enjeu clé pour assurer la pérennité des développements de technologies d'intelligence artificielle en France. En santé par exemple, la très grande majorité des technologies d'IA développées par des *start-up* créées en France l'ont été avec des données provenant d'autres pays, induisant un risque de délocalisation des activités dans ces pays et d'affaiblissement de l'écosystème français .

Principales recommandations

Des recommandations d'actions sont proposées dans plusieurs domaines. Toutes les recommandations proposées favorisent la création ou l'accès à des données en flux (indispensable à la mise au point d'algorithmes d'apprentissages performants) et/ou favorisent également le développement d'expertises nationales en Intelligence artificielle en développant les lieux d'expérimentation de technologies d'IA.

Créer des zones d'expérimentation à l'échelle d'un hôpital ou d'un réseau de soins

Cette recommandation correspond à créer plusieurs zones d'expérimentation accessible à tous les acteurs innovants, notamment en intelligence artificielle. Chaque zone d'expérimentation permettrait d'accéder à un cadre administratif défini, une infrastructure communicante, des données et des acteurs disponibles pour tester des innovations.

Un premier type d'expérimentation pourrait consister à transformer un ou plusieurs hôpitaux existants en centre d'expérimentation de technologies numériques dont l'intelligence artificielle. Un investissement en infrastructure serait réalisé dans l'hôpital pour rendre accessible l'ensemble des données médico-économiques aux sociétés voulant innover. Chaque service de l'hôpital devra accepter de tester des innovations pendant une période de plusieurs mois. La dimension légale serait également totalement prise en compte par l'hôpital expérimental pour éviter aux acteurs innovants un travail lourd et fastidieux de mise en conformité.

Un second type d'expérimentation correspond à rendre accessible l'ensemble des données des professionnels de santé à l'échelle d'une ville de plusieurs milliers d'habitants : pharmacie, médecine générale, structures de soins, hospitalisation à domicile, etc. L'objectif serait le même, permettre à des acteurs innovants d'expérimenter facilement et rapidement des technologies d'intelligence artificielle. Les investissements seraient de même nature que pour l'hôpital (infrastructure communicante entre les acteurs, données, cadre légal d'expérimentation, etc.).

Ces expérimentations nécessiteraient plusieurs types d'investissements de nature différente :

- La définition d'un cadre administratif et légal particulier et standardisé permettant d'expérimenter sur le lieu.
- Un investissement en infrastructure numérique (capteurs, réseaux, Datacenter) pour permettre à l'ensemble des équipements de santé de fournir des données numériques sur leur usage, les mesures cliniques et biologiques qu'ils permettent, les statuts de fonctionnement, etc.
- La fourniture d'un système d'information capable de collecter ces données pour les rendre disponibles.
- Une organisation dédiée à l'expérimentation pour accueillir tous les acteurs souhaitant innover.

Cette recommandation vise à créer les conditions permettant de tester des innovations en IA avec des délais de mise en œuvre les plus réduits possible.

Créer un cadre légal d'expérimentation à partir de données

La recommandation correspond à créer un cadre légal rendant accessible plus rapidement et facilement des données des structures de santé pour expérimenter des services basés sur leur usage. La législation permettrait aux structures de soins en France de fournir rapidement leurs données aux chercheurs ou aux acteurs innovants qui en feraient la demande.

Le cadre pourrait être basé sur la création d'une liste de lieux d'expérimentations disponibles pour tous les acteurs souhaitant innover. C'est un mécanisme qui existe déjà pour la gestion de la liste des hébergeurs de données de santé par exemple. La gestion de cette liste est l'une des missions de l'Agence Française pour la santé numérique (ASIPSANTE). Une mission supplémentaire pourrait être confiée à cette agence pour créer et maintenir une liste des acteurs pouvant utiliser des données cliniques ou biologiques dans le cadre d'expérimentations.

Cette initiative pourrait s'inscrire dans le cadre de France Expérimentation créée en 2016 qui vise à simplifier le processus administratif pour les acteurs souhaitant innover. Elle doit être l'une des missions prioritaires du « Health Data Hub » annoncé par le ministère de la Santé en juin 2018. Les données seraient fournies à titre expérimental, pour une durée limitée par exemple.

Développer les challenges de données en santé

Cette recommandation correspond à développer les challenges de données en santé sur le même modèle que le challenge Epidemium en cancérologie. Un challenge en santé permet de mesurer l'évolution de la performance des algorithmes de diagnostic ou de suivi thérapeutique dans le temps.

Un nouveau challenge pourrait être structuré par aire thérapeutique (diabète, maladies cardiovasculaires, maladies du foie, etc.) et confié à l'institut référent en santé dans chaque domaine comme l'Institut du cerveau et de la moelle épinière.

En termes d'investissement important, chaque challenge nécessiterait la création d'un Dataset multidimensionnel (physiologie, environnement, génétique, etc.) pour tester les algorithmes développés. Comme pour Epidemium, l'organisation portant le challenge regrouperait laboratoires, structures de recherche et innovateurs.

Créer un marché protégé pour les *start-up* en santé

La structure de remboursement des soins en France ne favorise pas le développement de services payants indispensables au financement des *start-up* dans la durée.

Cette recommandation correspond à créer une législation motivant les structures de soins à réaliser une partie de leurs achats technologiques auprès de structures européennes de petite taille. L'objectif est de créer un environnement de marché propice à l'émergence de sociétés viables économiquement. Ce type d'initiative a déjà été mis en place par le Conseil régional d'Île-de-France en mars 2017³¹. Cette initiative, de type « small business act », impose d'investir 2 % de la commande publique francilienne dans des achats auprès de *start-up innovantes*.

Un tel dispositif pourrait être mis en place pour le secteur de la santé de manière à créer un débouché pour les *start-up*.

Créer des zones d'expérimentation de véhicules autonomes à l'échelle d'une ville ou d'une situation de transport à risque (autoroute, zone montagneuse)

Cette recommandation correspond à créer plusieurs zones d'expérimentation accessible à tous les acteurs innovants dans le domaine de la conduite autonome.

Une première zone d'expérimentation correspondrait à une ville d'au moins 50 000 habitants, représentative de situations de conduite en milieu urbain, mais avec une taille qui limiterait le besoin d'investissement en infrastructure communicante.

Une seconde zone d'expérimentation correspondrait à une situation de conduite particulière :

- Une autoroute (plus de 100 km) pour les situations de conduite rapide.
- Une zone montagneuse pour les situations de conduite sur routes sinueuses, à forte déclinaison et/ou en conditions hivernales.

Ce projet nécessitera plusieurs types d'investissements de nature différente :

- La définition d'un cadre administratif et légal particulier et standardisé permettant d'expérimenter sur les lieux.
- Un investissement en infrastructure numérique (capteurs, réseaux, Datacenter) pour permettre à l'ensemble des équipements de transport (signalisation, mesure du trafic, etc.) de fournir des données numériques sur leur usage, leur statut de fonctionnement, etc.
- La fourniture d'un système d'information capable de collecter ces données pour les rendre disponibles.
- Une organisation dédiée à l'expérimentation pour accueillir tous les acteurs souhaitant innover

Cette recommandation nécessitera de créer les conditions permettant de tester rapidement une innovation dans le domaine de l'intelligence artificielle.

³¹ <https://www.iledefrance.fr/toutes-les-actualites/un-small-business-act-faciliter-l-acces-tpe-pme-aux-marches-publics-regionaux>

Amplifier les zones d'expérimentation à l'échelle d'un écoquartier

Cette recommandation correspond à créer plusieurs zones d'expérimentation correspondant à la taille d'un écoquartier de plusieurs milliers habitants. Cet écoquartier comprendra l'ensemble des technologies de transition énergétique : *smart city*, production et consommation intermittente, véhicule électrique, transport en commun électrique, etc.

Cette recommandation pourra être reliée à un des programmes Smart Energy en cours en vue de contribuer à amplifier par l'IA, des expérimentations déjà lancées :

- FLEXGRID (Conseil régional de Provence-Alpes-Côte d'Azur),
- SMILE (Conseil régional de Bretagne, en lien avec les Pays de la Loire)
- YOU & GRID (métropole européenne de Lille, en lien avec les Hauts-de-France).

Cela nécessitera plusieurs types d'investissements de nature différente :

- La définition d'un cadre administratif et légal particulier et standardisé permettant d'expérimenter sur les lieux.
- Un investissement en infrastructure numérique (capteurs, réseaux, Datacenter) pour permettre à l'ensemble des équipements de fournir des données numériques sur leur usage, leurs statuts de fonctionnement, etc.
- La fourniture d'un système d'information capable de collecter ces données pour les rendre disponibles.
- Une organisation dédiée à l'expérimentation pour accueillir tous les acteurs souhaitant innover.

Cette recommandation nécessite de créer les conditions permettant de tester rapidement une innovation dans le domaine de l'intelligence artificielle.

Généraliser et normaliser le recueil de données de maintenance

Cette recommandation vise à inciter les acteurs industriels à accélérer la mise d'un cadre global (infrastructures numériques, compétences, processus), favorisant le développement des activités de maintenance prédictive.

Alors que la refonte des systèmes de pilotage des usines du futur intégrant des algorithmes d'IA sera probablement dominée par les grands acteurs actuels (SIEMENS, BOSCH...), les innovations qui touchent à la maintenance sont plus variées (ex. : applications non invasives de *machine learning* pour capter et analyser en temps réel des données des machines). Elles offrent un potentiel plus fort pour des *start-up* et des PME nationales. Certaines connaissent déjà un développement prometteur (ex. : TELMEPLUS, CARTESIAM, SCORTEO...).

Les activités de maintenance sont profondément transformées par le déploiement progressif et conjoint de l'Internet industriel (Internet des objets appliqué au domaine industriel), de plateformes de Big Data et d'applications d'IA. Grâce à ces mutations, des domaines d'activité, des processus et des données en silo vont pouvoir être rapprochés et des équipements jusque-là faiblement suivis, pourront l'être davantage.

L'enjeu est de mieux comprendre le comportement des équipements et anticiper les défaillances éventuelles, de disposer d'informations en temps réel permettant un meilleur pilotage humain, économique et financier des opérations de maintenance (gestion des interventions, gestion des stocks, relation client...).

La puissance publique pourrait :

- Favoriser les échanges entre les acteurs pour renforcer les normes portant sur les activités de maintenance, intégrant plus systématiquement l'instrumentation par l'Internet industriel et l'analyse prédictive. Les bénéfices à mettre en avant auprès des industriels pour susciter leur adhésion sur cette évolution sont nombreux : gains économiques liés à une optimisation des opérations de maintenance, réduction des risques industriels et des coûts associés en sécurité, etc.
- Financer, sur la base d'appels à projets, les premiers pilotes.

Stimuler la production de données environnementales chez les industriels

Cette recommandation est complémentaire à celle présentée dans les recommandations pour le secteur Énergie-Environnement, intitulées : « renforcer les normes et règlements sur l'instrumentation et la collecte de données portant sur la qualité de l'air ». Elle vise à enrichir sensiblement la quantité, la variété, le degré de finesse et la fraîcheur des données relatives à la pollution industrielle.

L'industrie manufacturière est responsable du rejet d'un grand nombre de polluants atmosphériques. Elle est notamment le premier émetteur de zinc, de sélénium, de plomb, de mercure, de chrome, de cadmium, d'arsenic, de SF6 (hexafluorure de soufre), de PFC (perfluorocarbures) et de composés organiques volatils dans l'air.

Plusieurs maladies graves (cancers, troubles du développement chez le fœtus, affections ORL complexes, etc.) sont corrélées à ces émissions avec des risques à la fois pour les salariés et les riverains.

Aujourd'hui, les pouvoirs publics émettent des lois et règlements *a priori*, vérifiés par des audits ponctuels.

La recommandation correspond à favoriser, dans les usines, le déploiement plus systématique d'instruments (capteurs, robots) de collecte de données environnementales sur les rejets dans l'atmosphère à des fins de traitement par des algorithmes d'intelligence artificielle.

La recommandation s'appuierait sur deux leviers : la normalisation et le cofinancement de pilotes d'expérimentation.

- Faire effet de levier sur la normalisation.
 - Augmenter les mesures : promouvoir le monitoring par des incitations spécifiques.
 - Revisiter/réviser les normes existantes avec une vision digitalisation et IA.
- Contribuer au financement des expérimentations.
 - Concevoir et mettre en œuvre un crédit d'impôt pour investissement sur le *monitoring* sur des usines existantes.
 - Lancer des appels à projets.

Former rapidement des techniciens, des ingénieurs et des décideurs de l'industrie aux cas d'usages et aux techniques probabilistes de l'IA

Cette recommandation vise à acculturer et à former un grand nombre de dirigeants, ingénieurs et techniciens du secteur industriel, en particulier, dans sa composante Petites et Moyennes Industries (PMI) aux apports potentiels de l'intelligence artificielle. Cette formation devrait porter sur :

- Le partage d'expérience autour de cas d'usage concrets : visites de sites ayant déployé des applications IA, plateforme de partage de cas d'usage et espaces de discussion.
- Les méthodes probabilistes, les acteurs et les outils informatiques disponibles pour déployer des usages.

Pour être efficaces et ciblées, ces démarches de formation pourraient être :

- Promues par les pouvoirs publics.
- Mises en œuvre en partenariat entre l'Alliance Industrie du Futur, les différents syndicats industriels et des acteurs spécialisés du numérique.
- Déployées sur le terrain avec l'aide des collectivités locales et des chambres de commerce et d'industrie.

Créer un Datacenter de projets en data science en s'appuyant sur les expertises d'OVH et de TERALAB

Selon l'étude réalisée par l'association ALLISTENE regroupant notamment le CEA, l'IMT, le CNRS et l'Inria, la France ne dispose pas d'infrastructure de calcul pour réaliser des projets en IA permettant de répondre aux défis actuels : gestion des *SmartGrids*, changement climatique, mise au point du véhicule autonome, etc.

L'une des plus grandes infrastructures publiques, le Centre régional Informatique et d'Applications numériques de Normandie (CRIANN) dispose d'une puissance de 600 TFLOPS (CPU), 170 TFLOPS (GPU) et 2,5 Po de capacité de stockage. Le système de transcription de la parole DEEPSPEECH de BAIDU (un réseau de neurones comportant entre 18 et 100 millions de neurones selon le modèle) a nécessité l'entraînement de plus de 500 modèles, chaque modèle mobilisant 50 TFLOPS sur des processeurs GPU. Ramené aux capacités du CRIANN, l'entraînement aurait mobilisé l'ensemble des ressources pendant plus d'un an, ce qui n'est pas réaliste.

Pour répondre aux besoins d'entraînement de très grands réseaux de neurones, nous recommandons de créer une infrastructure de data science *hyperscale*³² souveraine et capable d'adresser tous les problèmes d'IA. Le projet ne vise pas à créer une nouvelle infrastructure de données, mais à s'appuyer sur des initiatives et moyens déjà largement disponibles. Concrètement, le projet pourrait reposer sur un partenariat entre la structure de TERALAB et l'infrastructure de la société OVH, chacun disposant des expertises complémentaires nécessaires.

TERALAB, de par l'expérience acquise depuis six ans sur des projets de recherche en data science avec de grands acteurs privés, assurerait la maîtrise d'ouvrage du projet. OVH, en tant que seul acteur majeur européen d'infrastructure serait responsable de la maîtrise d'œuvre du projet.

L'infrastructure offrirait des services de stockage et de calcul haute performance, simples, sécurisés et abordables pour tous les projets en data science. À terme, une partie des moyens de calculs des universités pourrait basculer sur cette infrastructure.

Encourager des projets centrés sur les assistants intelligents pour le grand public

L'intelligence artificielle dans le traitement du langage naturel est l'un des axes les plus dynamiques aujourd'hui : recherche vocale, assistant du conseiller, assistant d'achat, etc.. Mais la grande majorité des innovations dans ce domaine vient des GAFAs qui adressent leurs propres besoins, ou de *start-up* qui adressent les besoins de grandes entreprises (B2B).

³² À très grande échelle de calcul

L'assistant intelligent réellement à l'usage des besoins quotidiens du grand public ne bénéficie pas de la même dynamique, compte tenu notamment de la plus grande difficulté pour trouver un modèle économique rentable. Un tel assistant serait utile, par exemple, pour :

- Faciliter la comparaison d'offres complexes comme des prêts ou des propositions contractuelles.
- Simplifier une démarche administrative.
- Aider à gérer sa santé.

Le développement d'assistants intelligents pour le grand public pourrait se manifester par des appels à projets ou le lancement de marchés publics à destination de start-up.

Créer des formations en ligne (MOOC) sur l'IA et les produits numériques

Compte tenu du caractère récent de l'IA, la majorité des acteurs interrogés manquent de profils adaptés à des projets d'intelligence artificielle en particulier, et en data science en général. Ces profils devraient posséder une triple expertise métier, mathématique et technologies digitales. Le besoin est urgent et nécessite un effort de formation rapide au moins pour travailler sur des problèmes simples de *machine learning* (classification, régression) à partir de données disponibles.

Un effort de formation pourrait être réalisé au travers de contenus de formation en ligne (MOOC) adressant les principaux sujets : problèmes de data science, bases statistiques, principales techniques de *machine learning*, outils et *framework*, passage à l'échelle des problèmes de data science, conception de produits numériques, etc.

Ces contenus de formation n'existent pas en nombre et en qualité suffisante en langue française. Les contenus de qualité proviennent de sociétés (Coursera) ou d'organismes universitaires américains (cours de Machine Learning d'Andrew Ng à Stanford University).

Un projet de MOOC sur les sujets clés d'IA pourrait être confié à un ensemble d'universités et de grandes écoles qui ont tous des initiatives de travail dans le domaine.

Les rapports Pipame déjà parus

- Industrie du futur : enjeux et perspectives pour la filière aéronautique, décembre 2018
- Marchés des objets connectés à destination du grand public, mai 2018
- Potentiel de développement de l'économie sociale et solidaire dans les quatre secteurs économiques, octobre 2017
- Les acteurs, l'offre et le marché de l'efficacité énergétique à destination de l'industrie, octobre 2017
- Perspectives de développement de la filière des drones civils à l'export, juin 2017
- Enjeux et perspectives des producteurs pour tiers de principes actifs et de médicaments, mars 2017
- L'avenir du marché de la téléassistance et des services associés, février 2017
- Futur de la Fabrication additive, janvier 2017
- Marché actuel et offre de la filière minérale de construction et évaluation à échéance de 2030, novembre 2016
- Enjeux et perspectives des industries du sport en France et à l'international, juin 2016
- Filières industrielles de la valorisation énergétique du sous-sol profond, mars 2016
- E-santé : faire émerger l'offre française en répondant aux besoins présents et futurs des acteurs de santé, février 2016
- Usages novateurs de la voiture et nouvelles mobilités, janvier 2016
- Enjeux et perspectives de la consommation collaborative, juillet 2015
- Mutations économiques du secteur de l'industrie des métaux non ferreux, mars 2015
- Les innovations technologiques, leviers de réduction du gaspillage dans le secteur agroalimentaire : enjeux pour les consommateurs et pour les entreprises, novembre 2014
- Benchmark européen sur les plateformes chimiques, quels sont les leviers pour améliorer la compétitivité des plateformes françaises ?, septembre 2014
- Relocalisations d'activités industrielles en France, décembre 2013
- Imagerie médicale du futur, octobre 2013
- Évolutions technologiques, mutations des services postaux et développement de services du futur, juillet 2013
- Chaînes logistiques multimodales dans l'économie verte, mars 2013
- Enjeux économiques des métaux stratégiques pour les filières automobiles et aéronautiques, mars 2013
- Étude sur la location de biens et services innovants : nouvelles offres, nouveaux opérateurs, nouveaux modèles économiques ?, janvier 2013
- Potentiel et perspectives de développement des plates-formes d'échanges interentreprises, janvier 2013
- Enjeux et perspectives des industries agroalimentaires face à la volatilité du prix des matières premières, octobre 2012
- Le développement industriel futur de la robotique personnelle et de service en France, avril 2012
- La gestion des actifs immatériels dans les industries culturelles et créatives, mars 2012
- Marché actuel des nouveaux produits issus du bois et évolutions à échéance 2020, février 2012
- M-tourisme, décembre 2011
- Étude prospective des bassins automobiles : Haute-Normandie, Lorraine et Franche-Comté, novembre 2011
- Dispositifs médicaux : diagnostic et potentialités de développement de la filière française dans la concurrence internationale, juin 2011
- Pratiques de logistique collaborative : quelles opportunités pour les PME/ETI ?, février 2011
- Maintenance et réparation aéronautiques : base de connaissances et évolution, juin 2010
- Mutations économiques dans le domaine automobile, avril 2010
- Mutations économiques dans le domaine de la chimie - volet compétences, février 2010
- Mutations économiques dans le domaine de la chimie, février 2010
- Réflexions prospectives autour des biomarqueurs, décembre 2009
- Mutations économiques pour les industries de la santé, novembre 2009
- Le commerce du futur, novembre 2009
- Dimension économique et industrielle des cartes à puces, novembre 2009

- L'impact des technologies de l'information sur la logistique, novembre 2009
- Logistique : compétences à développer dans les relations « donneur d'ordres-prestataire », novembre 2009
- Logistique et distribution urbaine, novembre 2009
- Logistique mutualisée : la filière « fruits et légumes » du marché d'intérêt national de Rungis, octobre 2009
- La logistique en France : indicateurs territoriaux, septembre 2009
- Étude de la chaîne de valeur dans l'industrie aéronautique, septembre 2009
- Diffusion des nouvelles technologies de l'énergie (NTE) dans le bâtiment, juin 2009

Crédits photographiques

Couverture (horizontalement de gauche à droite) : © MJ_Prototype - GettyImages ; © metamorworks - GettyImages ; © Sergey - Fotolia ;
© Ekkasit919 - GettyImages

Faire de la France un acteur majeur de l'intelligence artificielle (IA) est l'ambition de la stratégie nationale présentée en mars 2018 par le président de la République. La dynamique d'innovation et d'investissement est forte dans le domaine de l'IA qui est sujet à une concurrence mondiale intense. Acteurs privés et publics ont multiplié par dix leurs investissements au cours des cinq dernières années afin de maîtriser cette technologie stratégique, porteuse de promesses.

Conscients des très forts enjeux liés à l'IA, la Direction générale des entreprises (DGE), le Commissariat général à l'égalité des territoires (CGET) et TECH'IN France ont confié à Atawao Consulting l'étude « Intelligence Artificielle – État de l'art et perspectives pour la France ». Après un état de l'art des différentes technologies du domaine, l'étude propose une méthode de classification des secteurs potentiellement les plus transformés par l'essor de l'intelligence artificielle et établit une analyse macroscopique de son adoption par ceux-ci. Elle approfondit ensuite cette analyse pour quatre secteurs : Énergie et environnement, Transport et logistique, Santé et Industrie. Pour chacun de ces quatre secteurs, un bilan des opportunités générées par l'IA est établi et une stratégie cible à adopter est proposée.

Dans sa dernière partie, l'étude dessine une feuille de route ainsi que des recommandations sectorielles et transverses qui permettront à la France et à ses entreprises de relever les défis en matière d'intelligence artificielle.