



ACADÉMIE DES TECHNOLOGIES

POUR UN PROGRÈS RAISONNÉ, CHOISI ET PARTAGÉ

Renouveau de l'Intelligence artificielle et de l'apprentissage automatique

Commission technologies de l'information et de la communication

Rapport de l'Académie des technologies

Mars 2018

Académie des technologies
Grand Palais des Champs-Élysées - Porte C
Avenue Franklin D. Roosevelt - 75008 Paris
+33(0)1 53 85 44 44
secretariat@academie-technologies.fr
www.academie-technologies.fr

©Académie des technologies 2017

ISBN : 979-10-97579-07-4

**Renouveau
de l'Intelligence artificielle
et
de l'apprentissage automatique**

Résumé

L'Intelligence artificielle (IA) est un domaine de l'informatique, un ensemble de disciplines qui s'intéressent au raisonnement et à l'imitation des capacités humaines. Depuis longtemps l'IA a une large intersection avec l'apprentissage automatique, une approche qui permet d'extraire des connaissances à partir de données. Les progrès spectaculaires des réseaux neuronaux profonds, en particulier dans le domaine de la perception (comme la vision artificielle ou la reconnaissance de la parole), ont créé une très forte attention ces dernières années, mais le domaine de l'Intelligence artificielle ne se réduit pas aux réseaux neuronaux ou à l'apprentissage automatique.

L'Académie des technologies s'est intéressée à ce domaine parce qu'il est en grande partie tiré par la technologie et la pratique. En discutant avec de nombreux experts d'horizons variés, nous nous sommes concentrés sur les conditions nécessaires pour la mise en œuvre de ces approches et les conseils qu'il est possible de donner à ceux qui souhaitent passer à l'action, en particulier aux entreprises.

L'IA n'est pas une technique spécifique et facilement identifiable, mais un ensemble large de méthodes qui ont vocation à s'intégrer dans l'ensemble des méthodes et outils informatiques. La modélisation et la compétence métier jouent un rôle fondamental dans l'obtention des succès. En positif, du point de vue de la compétitivité de la France, cela signifie que ce n'est pas un « service » qui

serait fourni par un acteur dominant (par exemple, américain). En négatif, il y a un investissement important à faire pour construire les processus de traitement de données, les protocoles de mise au point et le savoir-faire de mise en œuvre.

L'IA n'est pas un but en soi, mais un moyen qui peut servir à des fins multiples et qui est donc destiné à pénétrer la majorité des pratiques et des environnements, dans l'entreprise comme dans la société civile. L'enjeu, stratégique et compétitif, est la maîtrise de ces méthodes par les acteurs de l'écosystème français. Les entreprises doivent développer leurs capacités de mise en œuvre en travaillant sur leurs infrastructures de données, sur leur environnement logiciel qui doit être ouvert au monde du logiciel libre et en favorisant le travail itératif en cycle court de petites équipes pluridisciplinaires. L'accès à des grandes capacités de calcul est essentiel car il détermine le cycle d'apprentissage : plus on a accès à des machines rapides et spécialisées, plus on peut faire d'itérations du processus complexe qui est décrit dans ce rapport.

Les domaines de l'Intelligence artificielle et de l'apprentissage comportent une dimension importante liée aux technologies et aux pratiques qui doivent recevoir l'attention et les moyens des pouvoirs publics. Il faut faire émerger des « centres d'essais et de validation » par domaine d'application, regroupant industriels et chercheurs au niveau français d'abord, puis au niveau européen.

Le développement de l'Intelligence artificielle s'appuie sur la disponibilité de larges corpus annotés de données. Il est essentiel de faire entrer les industriels européens dans une logique de « norme de processus utilisant les données » avec une « finalité d'usage » explicite (par opposition à la « finalité de collecte » de la loi Informatique et Liberté) et une forme de rescrit juridique associé à ces normes permettant d'échapper au droit commun du GDPR.

Les pouvoirs publics ont un rôle à jouer pour développer un écosystème favorable, par la mise à disposition de données publiques, par le développement de

RÉSUMÉ

la formation et, de façon plus globale, par une communication positive sur ces nouvelles technologies de l'information. La réussite de l'écosystème de l'Intelligence artificielle passe également par l'usage des applications et la collecte des données, ce qui suppose une relation pacifiée entre la société civile et ces nouvelles technologies.

Introduction

L'Intelligence artificielle (IA) est une discipline ancienne de plusieurs décennies qui est redevenue très visible depuis quelques années, sous la double impulsion de résultats spectaculaires — depuis la victoire de *Deep Blue* sur Gary Kasparov aux échecs en 1996 à celle d'un algorithme d'IA de deux chercheurs de *Carnegie Mellon University* au jeu de Poker en 2017¹, en passant par la victoire de Watson à *Jeopardy* en 2011 et celle d'Alpha Go sur le grand maître Lee Seedol au jeu de Go en 2016 — et d'avancées significatives sur des problèmes jugés difficiles comme la reconnaissance d'image ou la reconnaissance de la parole. L'apprentissage automatique (utilisé ici comme traduction de « *machine learning* ») est l'une des disciplines informatiques qui a le plus nourri les progrès de l'IA ces dernières années. La commission TIC (Technologies de l'information et de la communication) de l'Académie des technologies s'est emparée de ce thème en 2015 en se posant la question de savoir si nous assistions à une révolution, une accélération ou une évolution de ce domaine, et s'il fallait urgemment se poser la question des conséquences de ce renouveau.

////////////////////////////////////

1 <https://www.wired.com/2017/01/mystery-ai-just-crushed-best-human-players-poker/>

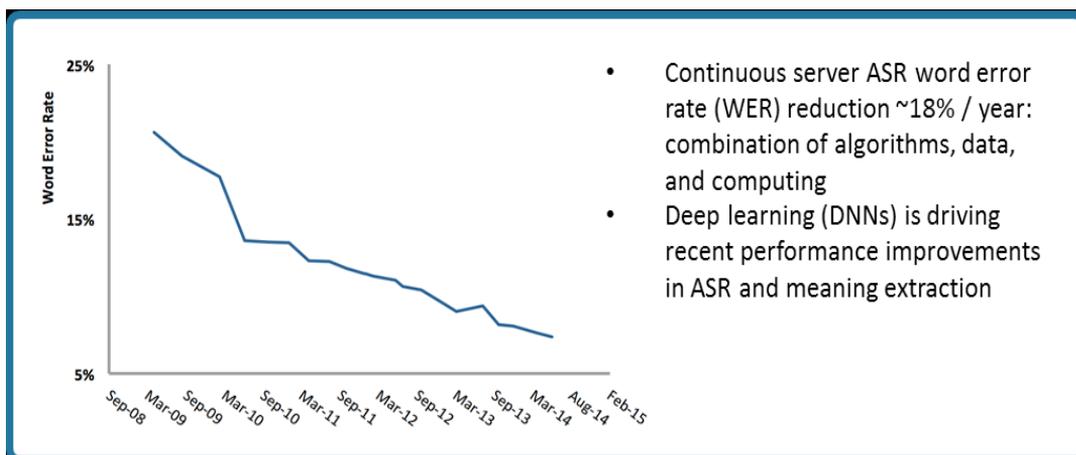


Figure 1 : les progrès en reconnaissance vocale

Il est clair qu'il se passe quelque chose de significatif depuis quelques années. La figure précédente montre les progrès spectaculaires obtenus en matière de reconnaissance vocale. Cela a permis à la machine de battre les meilleurs «*texteurs*» de la planète lors d'un concours organisé par BAIDU et l'université Stanford en 2016². Les progrès en reconnaissance d'image sont également spectaculaires. Le problème est ancien et nous avons pu constater des progrès très lents de 1990 à 2010, avec des niveaux de performance très sensiblement inférieurs à l'expertise humaine. Puis les choses ont rapidement changé et grâce au *Deep Learning* (que nous allons décrire dans la section suivante), les machines ont maintenant de meilleures performances que les humains³. Cette capacité s'étend à la reconnaissance des individus sur des images, ce que nous voyons apparaître dans notre vie quotidienne, dans nos smartphones ou sur des caméras connectées. Nous allons citer d'autres exemples dans le corps de ce rapport, mais on pense tout de suite à la performance de la voiture autonome de

2 Lire l'article : *Smartphone speech recognition can write text messages three times faster than human typing* <http://news.stanford.edu/2016/08/24/stanford-study-speech-recognition-faster-texting/>

3 Lire par exemple : http://www.eetimes.com/document.asp?doc_id=1325712. Cette performance n'est pas «*universelle*», elle correspond à un protocole de test précis. Ce qui est significatif, c'est que pendant plusieurs décennies, la vision artificielle était loin de la vision humaine, ce qui n'est plus le cas.

Google comme représentative de cette accélération des capacités de perception et d'analyse. Un autre indicateur de l'effervescence autour de l'Intelligence artificielle est l'accélération des investissements, qu'il s'agisse des grands acteurs ou des fonds de capital-risque — comme illustré par la figure 2. Google, Apple et Microsoft investissent chacun plusieurs milliards depuis plusieurs années⁴.

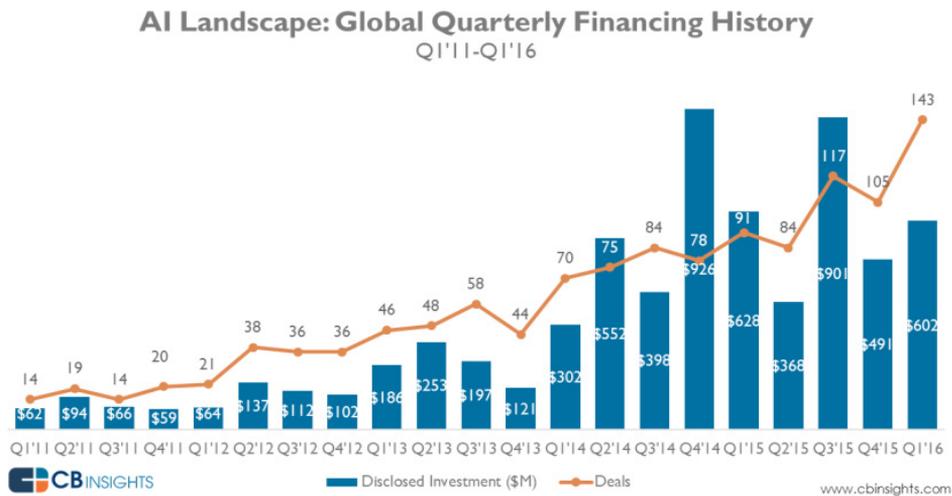


Figure 2 : l'intensification des investissements en Intelligence artificielle

Parce que le thème de l'Intelligence artificielle est très largement couvert⁵, la commission s'est concentrée sur que nous pouvions apporter au travers du prisme pluridisciplinaire technologie/mise en œuvre propre de l'Académie des technologies. Au-delà des progrès scientifiques, les technologies et leur mise en œuvre jouent un rôle fondamental dans le déploiement de l'Intelligence artificielle. Après avoir écouté de nombreux experts, il nous est apparu que, même

////////////////////

- 4 L'apprentissage est considéré comme la priorité numéro un par Satya Nadella, le CEO de Microsoft, pour un montant total d'investissement qui dépasse 10 milliards sur cinq ans.
- 5 Ce même raisonnement a été appliqué au précédent rapport de la commission TIC [1] — la commission cherche ici aussi à mettre en valeur les messages de nos experts pour les décideurs politiques et industriels, qui ne sont pas forcément suffisamment visibles dans la littérature « courante » sur le sujet.

s'il y a un consensus sur la plupart des recommandations issues des rapports déjà publiés, un certain nombre de questions méritent une réponse :

- savoir distinguer ce qui relève de l'accélération et ce qui constitue une rupture ;
- comprendre ce qui est rendu possible et éviter l'inflation des promesses qui risque assurément de produire des désillusions ;
- comprendre ce qui est nécessaire pour profiter de ce renouveau, en particulier pour les entreprises qui sont sommées par leur environnement d'adopter ces « nouvelles technologies ».

Il existe un florilège de citations qui expriment que nous sommes entrés dans une nouvelle ère, que l'Intelligence artificielle va bouleverser nos usages, nos marchés et nos entreprises. Il n'est plus possible de parler de stratégie ou de prospective sans y mettre l'IA en position centrale. La thèse que nous allons défendre dans ce rapport, à la suite de l'audition de multiples experts, est que l'IA est un outil, pas une fin en soi, mais que **la maîtrise de cet outil est bien un enjeu stratégique**. Cette maîtrise doit s'appuyer sur des compétences dont il faut favoriser l'émergence, ce qui signifie que cela prend du temps et qu'il n'y a pas de relation directe de cause à effet, donc que le volontarisme n'est pas suffisant. Ce rapport cherche en premier lieu à donner des clés pour favoriser la maîtrise des technologies autour de l'Intelligence artificielle et de l'apprentissage machine, dont on verra qu'elles sont nombreuses. Nous allons montrer qu'il y a une accélération des technologies et des performances dans le domaine de l'IA. Cette accélération ne doit pas non plus être surinterprétée : il reste beaucoup de défis à relever. Si l'on projette cette accélération sur la décennie à venir, il va y avoir une profonde transformation de nos usages, de nos entreprises, de la carte des pouvoirs économiques.

Le présent rapport est construit selon le plan suivant :

Le chapitre I propose un bref état de l'art pédagogique de ce qu'est l'Intelligence artificielle et de ce que sont les principales applications. Il existe de multiples rapports, livres et ressources en ligne qui abordent ce sujet, notre objectif est

donc simplement de permettre à l'ensemble des lecteurs de comprendre les propos qui vont suivre. Nous allons brièvement évoquer les différentes approches informatiques qui participent de l'Intelligence artificielle. Il s'agira de comprendre grossièrement comment elles fonctionnent, quels types d'applications utilisent ces méthodes aujourd'hui et pour quels problèmes. Cette section proposera également une synthèse des recommandations les plus fréquemment proposées, de ce qui fait consensus aujourd'hui au sujet de l'Intelligence artificielle. Même si le reste du rapport cherche à développer des positions plus originales, notre travail s'inscrit dans la continuité du rapport publié par France IA en mars 2016.

Le chapitre II cherche modestement à répondre à la question difficile « quelle IA pour quel problème ? ». La bulle d'intérêt médiatique sur les progrès récents du *Deep Learning* a tendance à créer une confusion contre laquelle la plupart de nos experts nous ont mis en garde. Même si la richesse du patrimoine de méthodes de l'IA est soulignée dans les rapports précédents, il nous a semblé intéressant de proposer un guide simplifié d'utilisation de la panoplie des méthodes disponibles, ce qui permet à la fois de préciser ce qui semble réaliste aujourd'hui et comment s'y prendre.

Le chapitre III est construit de façon symétrique comme une mise en garde contre les attentes irréalistes. Elles sont de trois types : l'exagération de ce que l'on sait faire aujourd'hui (ce que les Américains appellent le « hype »), l'application au mauvais type de problème (il devient courant de penser que tout est prévisible et possible de façon indépendante de la technologie dans ce monde du « big data » et de l'Intelligence artificielle, mais ce n'est bien sûr pas vrai) et les discours prospectifs simplistes qui font penser que tout va devenir possible, très vite.

Le chapitre IV est plus directement écrit pour les entreprises et traite des conditions favorables pour faire émerger la pratique de l'Intelligence

artificielle. Il y a une continuité entre ce rapport et le rapport précédent⁶ puisque la première condition de succès est la disponibilité de données qualifiées. Une des principales recommandations de nos experts est qu'il n'existe pas, pour l'instant, d'approche unique ou universelle, mais plutôt de multiples formes d'intelligences artificielles, adaptées à chaque domaine. Il convient donc pour les entreprises de faire émerger les pratiques et les compétences en adéquation avec les problèmes qui leur sont propres.

Le chapitre V contient quelques recommandations issues des entretiens avec les experts et des réflexions de la commission TIC.

/////////
6 Cf. rapport précédemment cité sur le « Big Data » de 2015 [1], dans sa 3^e partie qui traite de « data as code »

Chapitre I

État de l'art

COMPRENDRE L'INTELLIGENCE ARTIFICIELLE

Les multiples formes de l'IA

La définition intentionnelle précise de l'Intelligence artificielle n'est pas une chose simple et si l'on prend une approche large, elle couvre une très grande partie de l'informatique. L'ambiguïté commence par le mot « intelligence ». Ce terme peut être pris en référence à l'intelligence humaine, faisant de l'IA la science de l'imitation du raisonnement humain, ou dans une acceptation performative, celle de prendre des bonnes décisions en fonction d'une reconnaissance efficace d'un environnement. Mais on pourrait arguer que le fait de prendre des bonnes décisions à partir de traitements de données est le rôle d'une grande partie de la discipline informatique. Dès la naissance de l'Intelligence artificielle, qu'il est habituel de faire remonter à la conférence de Dartmouth en 1956⁷, on trouve deux aspirations de l'IA qui sont étroitement mêlées, celle de l'imitation de l'homme dans la lignée de Turing et celle du « résolveur de problème universel » de John McCarthy. L'appréciation de l'intelligence est liée à la surprise et au plaisir de

7 Organisée au Dartmouth college par Marvin Minsky, John McCarthy et deux scientifiques seniors : Claude Shannon et Nathan Rochester d'IBM.

la découverte, c'est pour cela que le périmètre de l'Intelligence artificielle est vu comme mouvant ; les choses qui deviennent exécutées automatiquement et régulièrement par la machine cessent de nous sembler relever du domaine de l'intelligence. Wikipédia cite l'exemple de la reconnaissance optique des caractères qui n'est plus vue comme une forme d'Intelligence artificielle mais comme une simple technique d'automatisation.

Il est plus facile de décrire l'Intelligence artificielle comme un ensemble d'approches, ayant chacune des objectifs plus précis que le « raisonnement intelligent ». La plupart des ouvrages sur l'Intelligence artificielle la définissent comme un ensemble de disciplines. Cette approche en extension est plus pragmatique, et en fait assez stable dans le temps. Dans le rapport de la *Maison Blanche* que nous citerons plus tard (2.3.1), c'est la taxonomie proposée par Franck Chen qui est utilisée (Figure 3) :

- perception, un domaine très vaste contenant la reconnaissance des images, des formes, et des sons (et de toutes formes de signaux produits par des « senseurs »). La reconnaissance des images est le domaine le plus important (on parle de *machine vision*) à cause de ses applications et de par les progrès spectaculaires de ces dernières années, cités en introduction, et liés au *deep learning* ;
 - traitement du langage naturel, sous sa forme écrite ou orale (ce qui implique de le coupler avec la perception de la parole) ;
 - planning et navigation, qui pourraient être étendus aux domaines de la résolution de problèmes formalisés dans l'univers de la recherche opérationnelle ;
 - représentation de connaissances, qu'il s'agisse de les manipuler, de les rechercher ou de les combiner. Ce domaine a fortement évolué avec le développement des mégadonnées (*Big Data*) ;
 - raisonnement logique, liée à la forme symbolique de l'Intelligence artificielle, rendue célèbre dans les années quatre-vingt par les « systèmes experts ».
- Ces différentes formes correspondent plus ou moins à des objectifs différents : la reconnaissance de situations, la mémorisation et l'apprentissage, le raison-

nement, la capacité d'expression et la décision (ou l'aide à la décision). Toutes ces approches contribuent à l'objectif de produire une « forme d'intelligence artificielle », cependant aucune n'est suffisante à elle seule pour se rapprocher de l'objectif de l'imitation de l'intelligence humaine. Il existe plusieurs variantes de la taxonomie de la figure 3, mais elles se ressemblent toutes. Le livre de cours le plus cité sur l'IA est celui de Russel & Norvig : « *Artificial Intelligence : A Modern Approach* » [26] ; il est organisé autour d'une table des matières similaire (résolution de problème, raisonnement sur les connaissances, raisonnement incertain, apprentissage, traitement du langage naturel et perception). Comme le remarque Yves Demazeau, la classification de l'IJCAI en 1983 n'est pas très différente de celle de l'AFIA aujourd'hui⁸.

Nous sommes tous, tel Monsieur Jourdain, des utilisateurs de l'IA dans notre vie de tous les jours, sans forcément le savoir [8]. Comme nous l'a expliqué Norm Judah, nous employons tous les jours sur nos ordinateurs des outils qui utilisent les méthodes d'apprentissage pour nous rendre des services « intelligents ». Les exemples les plus évidents sont les correcteurs orthographiques, les filtres antispam, les moteurs de recherche et les assistants⁹. De façon moins directement visible, les algorithmes de détection de fraude ou de recommandation sont des applications directes des techniques d'Intelligence artificielle. Les objets connectés, qu'ils s'agissent de voitures ou de smartphones, bénéficient déjà de capacités de perceptions avancées qui s'appuient sur les algorithmes que nous allons décrire plus loin. Le capteur Kinect de Microsoft utilise des techniques de *deep learning* pour l'analyse des mouvements à partir de la caméra.

8 AFIA : Association Française Intelligence artificielle, IJCAI : International Joint Conference on Artificial Intelligence - La même remarque s'applique au livre « *Exploring Artificial Intelligence – Survey Talks from the National Conferences on Artificial Intelligence* » paru en 1986

9 Nous y reviendrons plus loin, mais on peut se référer à l'article suivant du Web : <https://www.techemergence.com/everyday-examples-of-ai/>

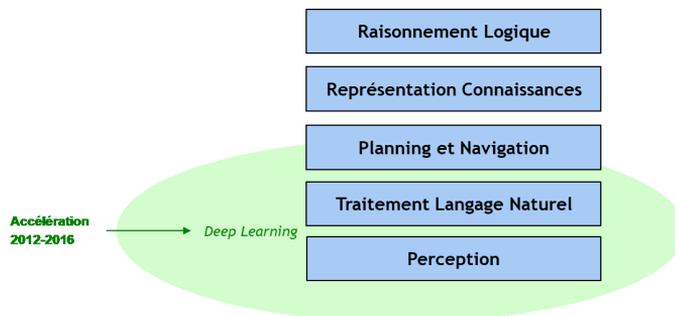


Figure 3 : Différentes formes d'Intelligence artificielle

Lorsque l'Intelligence artificielle cherche à construire un « résolveur de problème général », elle devient hybride et combine plusieurs de ces techniques.

La plupart des applications de l'Intelligence artificielle correspondent à un but précis et utilisent une famille de techniques de la taxonomie de la Figure 3. Mais lorsque l'objectif est de résoudre une classe plus large de problème, il devient nécessaire de mélanger plusieurs approches. La Figure 4 est extraite d'une présentation du *Todai Robot* dont l'objectif est de pouvoir réussir le concours d'entrée à l'Université. C'est un objectif de recherche ambitieux — le concours d'entrée au Japon ressemble au baccalauréat : il s'agit de lire des textes et de rédiger des réponses, pas de remplir un QCM — avec un programme de recherche sur dix ans. Le robot est déjà capable de « réussir le bac et rentrer dans une université de deuxième rang ». Il utilise une large palette des méthodes évoquées plus haut.

Cette remarque s'applique également à *Watson*, la plate-forme cognitive d'IBM, rendue célèbre par son succès à *Jeopardy*, mais qui a également démontré une capacité à faciliter des diagnostics médicaux ou répondre à des questions difficiles à partir d'une base de connaissance¹⁰. *Watson* utilise un large

//////////
 10 Comme d'autres plates-formes, *Watson* a été victime d'abus de marketing et a donc été probablement « survendu », lire par exemple : <https://www.statnews.com/2017/09/05/watson-ibm-cancer/>. Il n'en reste pas moins que de nombreuses applications réussies utilisent *Watson*.

ensemble de méthodes de la famille de l'IA : traitement du langage naturel, représentation de connaissances et raisonnement déductif, apprentissage machine et réseaux neuronaux profonds. *Watson* cherche en premier lieu à résoudre des problèmes de classification, dans leur acceptation la plus générale. Les méthodes d'apprentissages et de réseaux neuronaux sont intégrées dans un ensemble de méthodes de traitement statistiques des données. Il existe d'autres plates-formes d'Intelligence artificielle ayant une vocation « généraliste », telles qu'*Einstein* de Salesforce¹¹, ou *Holmes*¹² de Wipro, qui sont également des exemples d'intégration de différents types d'outils de classification et prévision.

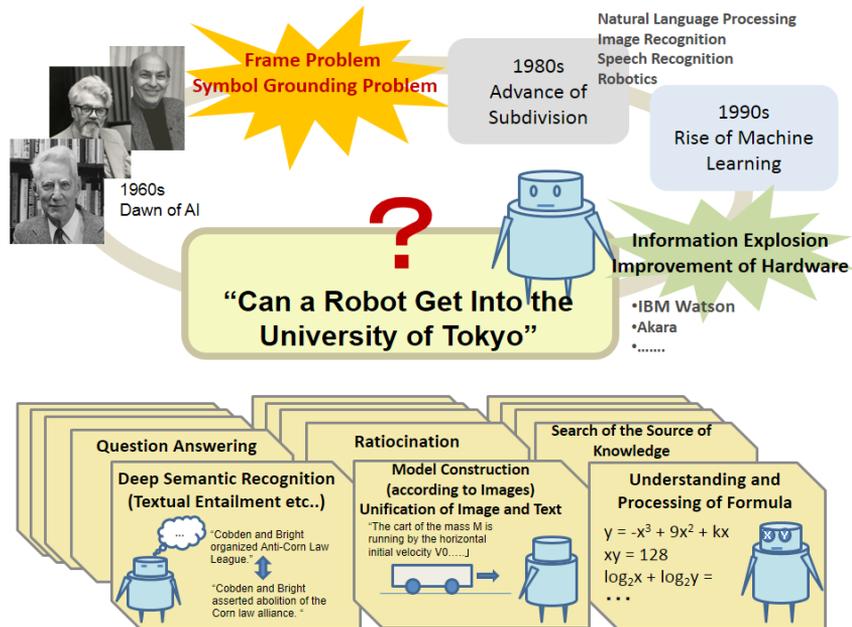


Figure 4 : Le « *Todai Robot* » de l'Université de Tokyo

- 11 *Einstein* est une plate-forme proposée par la société Salesforce dont l'objectif annoncé est « d'apporter l'intelligence artificielle pour tous ». La partie la plus intéressante est l'ambition de permettre à tous les utilisateurs de Salesforce de bénéficier d'une partie des avancées que nous allons citer dans ce rapport de façon assistée et semi-automatisée. Voir par exemple : <http://www.journaldunet.com/solutions/dsi/1183595-salesforce-einstein-intelligence-artificielle-3/>
- 12 *Holmes* (un clin d'œil à Watson, qui est l'acronyme de « Heuristics and Ontology-based Learning Machines and Experiential Systems ») est également une plate-forme d'IA hybride – cf. Chapitre III

Il est devenu courant de distinguer l'Intelligence artificielle étroite (ANI : *Artificial Narrow Intelligence*) et la version générale (AGI : *Artificial General Intelligence*). Il est clair que les applications d'aujourd'hui sont de type « étroit » : un problème bien défini, pas de véritable capacité à s'adapter aux situations imprévues et une absence de réflexion (raisonner sur soi-même) et d'autoévaluation¹³. En revanche, il ne faut pas prendre cette distinction de façon binaire. Le développement des nouvelles capacités d'intelligence artificielle dans les laboratoires de recherche tend précisément à ouvrir le domaine d'applications vers une plus grande généralité, à introduire des formes de créativité et d'adaptation aux aléas de l'environnement, voire à se doter d'une forme d'autonomie. Citons l'exemple d'*Einstein* qui est en fait une boîte à outils assez classique, mais dotée d'une capacité en amont à générer des modèles et à trouver par apprentissage les meilleures techniques de la boîte à outils.

Ce que l'IA sait faire aujourd'hui

Soulevons le couvercle de la boîte à outils

Il n'est pas possible dans ce rapport d'entrer dans le détail des modes de fonctionnement des différentes méthodes, mais nous allons évoquer quelques principes généraux, et fournir une bibliographie permettant d'aller plus loin. Une compréhension même grossière permet d'apprécier la complémentarité des différentes familles que nous avons présentées plus haut. *Nous prenons ici volontairement une approche large* en incluant des disciplines comme la recherche opérationnelle ou la démonstration de théorèmes, qui ne sont pas vues comme des branches de l'IA par les chercheurs de ces domaines, mais pour lesquelles il y a une forte proximité des objectifs généraux et une « perméabilité » (co-influence) des méthodes avec celles issues des communautés plus « historiques » de l'intelligence artificielle.

13 On parle également, de façon assez semblable, d'IA faible et forte. Pour une discussion plus approfondie et plusieurs références dont les excellents billets de Tim Urban, lire : <http://informationssystemsbiology.blogspot.fr/2015/06/strong-artificial-intelligence-is.html>

- Les méthodes déductives / symboliques utilisent des principes de logique pour déduire des faits nouveaux à partir de faits élémentaires ou vérifier des propriétés par déduction logique¹⁴. Durant les dernières décennies se sont développées des formes étendues de logique permettant d'étendre le champ d'application à des faits incertains (par exemple la logique floue qui assigne une véracité numérique aux faits) ou à des modes de déduction modaux (tenant compte de points de vue différents). La preuve automatique de propriétés est une application pratique de la démonstration de théorèmes. Elle a de nombreuses applications dans le domaine de la validation de propriétés de logiciels — par exemple en aéronautique (A380 ou A350) ou ferroviaire (Métro automatique type Ligne 14 à Paris) — ou de validation de protocoles, qu'il s'agisse de paiements ou de cybersécurité¹⁵.
- La recherche opérationnelle fournit une grande panoplie d'algorithmes pour les problèmes de décision et de planification. On y retrouve les méthodes classiques d'ordonnancement, d'optimisation de transport, d'allocation de ressource, etc. Lorsque les problèmes sont structurellement trop complexes ou hétérogènes, les plates-formes d'agents « intelligents » fournissent une méthode très générale de simulation, d'apprentissage et d'aide à la décision. Une telle plate-forme permet de mettre en réseau un très grand nombre d'« agents » qui sont programmés pour interagir avec une partie de leur environnement qui représente le problème à étudier¹⁶

////////////////////////////////////

- 14 Le développement de l'IA symbolique s'est accompagné de multiples méthodes d'exploration, de planification et de construction de stratégie de jeux. On parle parfois de GOFAI (*Good Old-Fashioned AI*) pour évoquer cette boîte à outils riche qui est encore utilisée en combinaison avec des méthodes plus modernes [19].
- 15 Elle a également permis, au-delà de la formalisation complète de preuves complexes en mathématique (le théorème des quatre couleurs ou le théorème de Feit-Thomson) de prouver des conjectures mathématiques ayant résisté des dizaines d'années aux mathématiciens humains non assistés par des machines : la première d'entre elles étant la conjecture de Robbins qui a tenu cinquante ans et a été résolue le prouveur EQP de Mc Cune (http://ftp.mcs.anl.gov/pub/tech_reports/reports/P642.pdf.)
- 16 Par exemple, REPAST est une plate-forme d'agents intelligents qui permet aujourd'hui de réaliser des simulations avec des millions d'agents ([https://en.wikipedia.org/wiki/Repast_\(modeling_toolkit\)](https://en.wikipedia.org/wiki/Repast_(modeling_toolkit)))

et pour collaborer en vue d'une résolution ou d'une caractérisation du problème. Chaque agent utilise le plus souvent des méthodes symboliques et déductives pour décrire son comportement, mais la multiplicité des agents permet de faire émerger des comportements riches et complexes¹⁷. L'approche par agents intelligents est à la frontière entre l'intelligence artificielle et l'étude des systèmes complexes.

- La boîte à outils contient également toutes les méthodes de classification des données issues des statistiques, qui sont les méthodes du *Big Data*. Ces méthodes permettent d'analyser et de classifier des données (par exemple sous forme de clusters ou d'arbre de décision), ou d'extrapoler des prévisions. Ces méthodes sont bien connues pour la plupart, on trouve une excellente introduction dans le livre *Data Science for Business* [9].
- Dans le domaine de la perception, la méthode qui domine les autres est clairement celle des réseaux neuronaux profonds et, en particulier, les réseaux convolutifs dans le domaine de la vision. Nous y revenons au chapitre suivant, compte tenu de l'importance du Deep Learning. La popularité de cette approche est telle que nous allons adopter le terme anglais dans la suite du rapport.
- Le traitement du langage naturel fait également appel au *Deep Learning*, par exemple dans les applications de traduction, pour lesquelles le *Deep Learning* domine toutes les autres approches, en particulier les approches linguistiques plus anciennes¹⁸. Tout comme pour la correction orthographique, les meilleurs algorithmes sont obtenus par apprentissage sur des corpus massifs de milliards d'occurrences. L'approche linguistique n'a pas disparu puisque les corpus lexicographiques et les ontologies sont les fon-

17 Par exemple, on peut citer la plate-forme de Cosmo Tech (<https://cosmotech.com/>) qui allie modélisation de systèmes complexes et agents intelligents pour construire une « intelligence augmentée ».

18 Comme le souligne un de nos experts, l'applicabilité est limitée par la disponibilité de vastes corpus annotés, mais l'approche *Deep Learning* gagne toutes les compétitions Kaggle qui portent sur des problèmes de traitement du langage naturel.

dations des outils de représentations de connaissances¹⁹ et des capacités « intelligentes » des nouveaux moteurs de recherche.

Le cas particulier des réseaux neuronaux profonds

Le *Deep Learning* est une méthode de classification / reconnaissance de « motifs » qui peuvent être très complexes, comme des images, des situations financières ou des parties de Go. Les réseaux neuronaux sont un assemblage en couche de grilles d'éléments qui reproduisent un schéma d'activation très simple, une abstraction grossière d'un neurone [7] [12]. Le « neurone » d'un réseau neuronal transmet un signal aux neurones sortants (le maillage du réseau est orienté) si la somme des signaux entrants dépasse un certain seuil. L'intérêt du concept de réseau neuronal vient des algorithmes de rétro propagation²⁰ qui permettent de calculer les paramètres de la grille de neurones à partir des signaux d'entrée (les données que l'on veut classer) et des signaux de sortie (le classement). Cette approche est fort ancienne (plusieurs décennies), mais elle a buté pendant de nombreuses années sur la complexité computationnelle (le temps de calcul prohibitif pour entraîner des millions de neurones, nécessaires à une bonne classification).

Le terme de « profond » vient du nombre de couches nécessaires à l'obtention de bonnes performances de classification, tandis que l'explosion que la discipline du *Deep Learning* a connue ces dernières années vient des progrès de la puissance de calcul et de la création de très vastes corpus de données d'entraînement. L'explosion de la vitesse de calcul est la conséquence de la Loi de Moore, et de l'utilisation de formes de matériels plus spécialisés, en

- ////////////////////
- 19 On parle de « Web sémantique » pour désigner les différentes approches qui permettent de regrouper les données du Web en fonction de leur sens, au-delà des graphes constitués des données et des pages du Web. Par exemple, les ontologies permettent de classer, de regrouper et de comparer des termes et des informations distinctes, mais semblables du point de vue sémantique.
- 20 La rétropropagation utilise une approche de gradient (propagation d'écarts), c'est donc une méthode essentiellement numérique demandant un très grand nombre de calculs simples.

particulier des GPU²¹. La création des corpus de données, telle que la base ImageNet qui contient plusieurs dizaines de millions d'images catégorisées, a joué un rôle fondamental dans le progrès de ces dernières années que nous avons rappelé en introduction. Le terme de convolutif s'applique à des réseaux dont les couches sont organisées pour permettre de détecter des motifs (par convolution d'un bloc), et pour progresser en degrés successifs d'abstraction. Les réseaux convolutifs ont des propriétés d'adaptation à l'échelle qui les rendent particulièrement adaptés à la vision²².

L'apprentissage d'un réseau multicouches est plus complexe puisque l'empilement des couches permet un encodage exponentiellement plus riche. Le nombre de couches nécessaires semble lié à la complexité structurelle du problème. Il y a eu plusieurs progrès décisifs au cours des dernières années, en premier lieu avec une approche d'apprentissage couche par couche, correspondant à une montée progressive en abstraction des formes à reconnaître. Depuis, il existe des méthodes qui font de la rétropropagation sur l'ensemble des couches, mais qui demandent beaucoup de savoir-faire et d'essais à cause de la complexité du « paysage d'optimisation » liés à la non-linéarité de la fonction d'activation des neurones²³. Plusieurs de nos experts ont employé le terme de « cuisine », non dans un sens péjoratif, mais parce que cet apprentissage est un art autant qu'une science. En effet, si les principes des réseaux neuronaux multicouches sont bien établis et si les codes sont facilement accessibles, il reste beaucoup de savoir-faire dans le « meta-paramétrage » de ces algorithmes. Pour ces experts, il faut considérer les réseaux neuronaux profonds comme une tech-

21 GPU : *Graphical Processor Unit*. Les GPUs ont des grandes capacités en calcul sur les nombres flottants et leur architecture, plus simple que celle des CPU, se prête bien à la parallélisation.

22 Le principe des réseaux convolutifs dépasse le cadre de ce rapport. Pour une introduction, lire : <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/> ou [12].

23 Les chercheurs de Microsoft Research ont utilisé un réseau à 152 couches en 2015 pour gagner la première place du challenge ILSVRC 2015 – il existe aujourd'hui des approches qui s'appuient sur un nombre encore supérieur de couches. http://image-net.org/challenges/talks/ilsvrc2015_deep_residual_learning_kaiminghe.pdf

nique d'apprentissage modulaire et combinable avec d'autres approches. De fait, le champ d'application est large, et dans de nombreux domaines, les DNN (*deep neural nets*) ont globalement remplacé les SVM (*support vector machines*) dans les concours d'algorithmes, alors que les SVM étaient en tête il y a dix ans.

D'autres techniques d'apprentissage

L'apprentissage par renforcement est une méthode générique qui peut se combiner avec la plupart des approches évoquées dans ce chapitre. C'est également une approche très ancienne, qui consiste à définir une « fonction objectif » qui représente la performance de l'algorithme, qui est ensuite utilisée pour optimiser les paramètres de l'algorithme (de décision, de reconnaissance, de classification) pour qu'il s'améliore de façon itérative. Par exemple, *AlphaGo* utilise à la fois les réseaux neuronaux profonds et l'apprentissage par renforcement. Cette technique est utilisée très souvent en robotique²⁴. L'apprentissage par renforcement est vu comme une approche similaire à l'apprentissage des bébés et des jeunes enfants. *Libratus*, le système d'intelligence artificielle qui a obtenu les résultats spectaculaires au Poker que nous avons mentionnés en introduction est construit autour de l'apprentissage par renforcement, sans utilisation de réseau neuronal²⁵.

Les réseaux neuronaux profonds ont démontré leur capacité « à faire mieux que les humains » sur de nombreux problèmes, mais ils ont souvent la propriété de se tromper de façon plus brutale (lorsque la situation d'entrée est « anormale » par rapport au corpus d'apprentissage). L'approche dite « *adversarial training* » consiste à utiliser un autre algorithme d'apprentissage pour construire des contre-exemples qui vont rendre progressivement l'algorithme initial plus

24 Voir, par exemple, le robot italien qui apprend à faire des crêpes (<http://www.semageek.com/le-robot-qui-fait-sauter-les-crepes-dans-une-poele-apprentissage>) ou encore le robot Yamaha qui conduit une moto (http://www.huffingtonpost.fr/2015/10/29/video-robot-moto-yamaha_n_8416900.html.)

25 Lire l'article de Wired *Inside Libratus, the Poker AI that out-bluffed the best humans* : <https://www.wired.com/2017/02/libratus/>

robuste. Comme cela a été dit, on peut combiner les réseaux neuronaux avec d'autres formes d'apprentissage, comme l'apprentissage par renforcement ou l'apprentissage par transfert. L'apprentissage par transfert²⁶ permet de modulariser l'apprentissage et de « transférer » ce qui a été appris dans un domaine vers un autre domaine.

Une autre forme d'apprentissage qui est ancienne, mais qui joue un rôle important dans les applications pratiques, en *data mining* ou dans les systèmes adaptatifs tels que l'antispam, est l'apprentissage bayésien [4]. Comme son nom l'indique il s'agit d'une méthode statistique issue de l'application du calcul de probabilité conditionnelle sous forme de réseau de dépendance *a priori*, qui est mis à jour continûment en fonction des observations. Il existe encore de nombreuses autres formes d'apprentissage, en particulier dans le monde des méthodes stochastiques. Par exemple, les algorithmes de « bandit »²⁷ sont largement utilisés dans les applications d'e-commerce du Web, mais on pourrait tout aussi bien parler des *hidden markov models*²⁸. Pour terminer ce bref tour d'horizon, citons également les méthodes d'apprentissage par induction dont l'objectif est de produire des règles (donc en complément d'une approche logique / déductive précédemment citée) à partir de bases de connaissance. Une des déceptions des trente dernières années attachées à ce type d'Intelligence artificielle est que ces méthodes progressent lentement, contrairement à la reconnaissance de « motifs » avec la *Deep Learning*, ce qui suggère peut-être que la reconnaissance de règles est plus difficile.

////////////////////////////////////

- 26 Pour une introduction au sujet, lire *Transfer Learning : réaliser de meilleures prédictions avec peu de données* : <http://www.lebigdata.fr/transfer-learning-analytique-0201>
- 27 Le nom vient de l'analogie avec des problèmes de stratégie de choix face à des « bandits manchots » dans les salles de jeu. Pour une introduction sur l'application à l'optimisation des publicités en ligne, lire : <https://blog.octo.com/online-machine-learning-application-a-la-publicite-sur-le-web/>
- 28 La recherche combinée de HMM et e-commerce sur Google donne des dizaines d'exemples d'applications, mais on peut commencer avec l'article suivant : <https://www.linkedin.com/pulse/marketing-attribution-challenge-hidden-markov-model-sudipt-roy-phd/>

Méthodes d'exploration « créatives »

On reproche souvent aux algorithmes « intelligents » leur absence de créativité, par opposition à l'intelligence humaine. Mais cette dimension d'exploration créative²⁹ existe déjà dans de nombreuses applications de synthèse de pièces mécaniques, de planification, de synthèse musicale, littéraire ou artistique. Le principe général est de décrire un espace à explorer — par exemple un ensemble de formes ou un ensemble de programmes — et d'utiliser des techniques classiques de « randomisation » pour explorer cet espace qui peut être extrêmement large. Les techniques les plus fréquentes sont l'approche Monte-Carlo — célèbre par son utilisation en finance — ou le parcours aléatoire de voisinages — souvent utilisé comme technique d'optimisation. L'utilisation des « variations aléatoires » (*randomization*) pour améliorer les performances d'algorithmes d'aide à la décision s'est accélérée depuis les années 2000 suite précisément à l'accélération des performances.

Les approches « génératives » se combinent facilement avec l'apprentissage par renforcement puisqu'il est naturel de combiner la description d'un espace à explorer avec une fonction d'évaluation. Elles sont également bien adaptées à l'utilisation de plates-formes multi-agents. Il existe de nombreuses techniques pour augmenter la « créativité » de ces méthodes consistant à utiliser plusieurs familles de fonctions d'évaluation dont certaines représentent une forme de « surprise ». Par exemple, dans son livre *The Future of the Mind* [20], Michio Kaku présente la surprise comme un « méta-paramètre » d'acquisition de connaissance et d'amélioration par renforcement de la capacité de prévision.

L'approche multi-agents permet également de faire cohabiter des points de vue et des modes de raisonnements différents, voire contradictoires, au sein d'une même simulation. Ceci conduit à une autre hybridation, celle de l'Intelligence

29 Il aurait été logique de parler ici de méthodes « génératives », mais ce mot est un « mot valise » qui a plusieurs sens techniques bien précis dans différentes branches de l'IA, donc il y avait un risque de confusion. Pourtant le fait de retrouver ce terme dans plusieurs disciplines illustre ce courant de l'IA qui mélange « génération » et « apprentissage »

artificielle et de la théorie des jeux, en particulier sous forme de la théorie des jeux évolutionnaires³⁰. La théorie des jeux évolutionnaires fournit un bon cadre pour combiner des méthodes de génération, de recherche d'équilibre entre plusieurs acteurs et d'optimisation des comportements, qui s'implémente facilement dans une approche multi-agents. La théorie des jeux apporte à l'Intelligence artificielle un certain nombre d'outils pour caractériser les types de coopération ou compétition entre joueurs. L'ordinateur a une plus grande facilité que l'être humain à calculer et évaluer des stratégies mixtes (combinaison de différentes stratégies avec différentes probabilités), ce qui peut expliquer la grande efficacité de l'IA dans des jeux qui incluent des aspects de hasard comme les jeux de cartes³¹. Comme nous l'a déclaré Norm Judah : « *Evolutionary algorithms and meta-algorithms (strategy for other algorithms) are on the rise* ». Ce type d'approche est particulièrement indiqué pour appliquer l'Intelligence artificielle à l'aide à la décision pour la stratégie d'entreprise³².

Exemples d'application de l'IA

IA « classique »

Même si les systèmes experts ont été très décriés³³, ils existent cependant dans de multiples applications industrielles. Ces applications utilisent des bases de règles et des moteurs qui fonctionnent par déduction / propagation. Parmi les milliers d'applications (de nombreuses sociétés de logiciel ont construit leur succès sur ces technologies) on peut citer la configuration, la maintenance et le

- ////////////////////////////////////
- 30 La théorie des jeux évolutionnaire s'intéresse à la construction d'équilibres par simulation des réactions intelligentes des acteurs (<https://ai.vub.ac.be/research/techniques/evolutionary-game-theory>). Pour une introduction, lire *Evolutionary Game Theory* de Jorgen Weinbull [17]
 - 31 L'utilisation de l'IA dans les jeux vidéo pour simuler un ou plusieurs adversaires intelligents est très ancienne. L'évolution des capacités de calcul permet aujourd'hui de construire des stratégies très sophistiquées pour des groupes importants d'acteurs virtuels.
 - 32 Voir par exemple *Game-Theoretical and Evolutionary Simulation: A Toolbox for Complex Enterprise Problems*, Y. Caseau, Proceedings of Complex and Dynamic Systems Management, 2012
 - 33 Les deux premiers « hivers » de l'intelligence artificielle dans les années 70 (cf. le « rapport Lighthill ») puis fin des années 80 correspondent à l'incapacité de tenir toutes les promesses qui avaient été faites trop vite.

dépannage. Dans le domaine de la configuration, les règles sont utilisées pour décrire des contraintes de dépendance et d'incompatibilité (pour cette raison, les règles sont souvent mélangées avec des contraintes). Dans le domaine du dépannage, les systèmes de règles décrivent des arbres de décisions qui aident l'opérateur à parvenir plus vite à un diagnostic.

Les moteurs de règles utilisant la logique floue ont également trouvé depuis une vingtaine d'années de nombreuses applications dans le domaine industriel du pilotage des objets : pilotage de l'autofocus des appareils photos, gestion des cycles de machine à laver, contrôle de la température avec des thermostats, pilotage des freins dans les premières générations d'ABS, etc. Ces approches sont facilement implémentées dans des microcontrôleurs que l'on retrouve dans une grande partie de l'industrie³⁴.

Dans le domaine « classique » de l'aide à la décision, la recherche opérationnelle est un domaine où la machine fait mieux que l'humain depuis longtemps. Même s'il ne s'agit pas à proprement parler d'intelligence artificielle, il existe une frontière poreuse entre les domaines, nous allons y revenir par la suite. En particulier, les techniques issues de l'IA comme les règles et les contraintes sont intégrées de façon hybride dans de nombreuses applications industrielles de planification, ordonnancement ou allocation de ressources pour intégrer des contraintes propres à l'environnement humain, ou pour rendre les applications d'aide à la décision plus adaptables³⁵.

Méthodes statistiques pour l'IA

Les méthodes statistiques de classification et détection de *patterns* jouent depuis longtemps un rôle très important dans nos vies, en particulier dans le

////////////////////////////////////

34 Voir par exemple : <https://www.slideshare.net/virajpatel77701/application-of-fuzzy-logic>

35 Nous reviendrons sur cette hybridation des algorithmes de recherche opérationnelle (par exemple la planification) et les algorithmes d'apprentissage statistique (par exemple pour la prévision). Elle trouve son application – par exemple – dans l'optimisation des « supply chains » : <https://channels.theinnovationenterprise.com/articles/how-supply-chains-are-adopting-ai>

domaine bancaire. Une des applications les plus significatives est la détection de fraude utilisée dans un très grand nombre d'industries. AMEX est célèbre pour avoir employé très tôt des techniques d'IA (au sens large) pour repérer des « *patterns* » frauduleux. Le « *credit scoring* » qui joue également un rôle essentiel dans la vie des citoyens, en particulier aux États-Unis, est fondé sur des méthodes statistiques. Si les premières générations (cf. FICO aux US) s'appuyaient sur des méthodes simples, les nouvelles générations s'appuient sur la puissance complète du *Big Data* et de la boîte à outils³⁶.

Ici aussi il est impossible de lister de façon complète tous les domaines des entreprises qui font maintenant appel aux méthodes statistiques de l'analyse de données. Par exemple, dans le domaine de la relation client, depuis l'acquisition, l'*upsell* et le *cross-sell* (comment vendre plus), à la rétention et la fraude, les départements *business intelligence* des services marketing ont appris à utiliser la boîte à outils. C'est d'autant plus facile que de nombreuses sociétés spécialisées, telle que la start-up DaVinci Labs en Corée³⁷, aident ces entreprises à choisir les meilleurs outils pour optimiser leur performance commerciale. La plate-forme *Einstein* de la société Salesforce, précédemment évoquée, permet d'utiliser l'intelligence artificielle pour optimiser tous les processus de la gestion des prospects et des clients. Cet exemple est doublement intéressant : d'une part parce que la partie sophistiquée de la plate-forme est l'autoadaptation au contexte de l'entreprise (moins « étroit » qu'on pourrait le penser) et, surtout, parce qu'il s'agit d'une véritable démocratisation de l'IA qui est déjà en marche.

L'analyse de texte, qui s'appuie sur les outils NLP, a également trouvé des multiples applications, telle que l'analyse systématique des échanges avec les clients, y compris par transcription automatique des appels téléphoniques (*call*

36 Ce qui n'est pas sans poser des problèmes éthiques, ce qui est justement souligné dans le livre de Cathy O'Neil *Weapons of Math Destruction* [2].

37 Start-up fondée par Sophie Eom et Olivier Duchenne, qui propose une remarquable plate-forme d'application de l'IA pour les entreprises, qui démontre la complémentarité des méthodes « IA et Big Data », cf. <https://solidware.io/en>

mining). Ces recherches de motifs dans les conversations s'appliquent bien sûr en cybersécurité. Plus généralement, la cybersécurité est un grand consommateur de méthodes d'intelligence artificielle, dans un jeu permanent du « chat et de la souris » entre les « détecteurs » et les « intrus ». Par exemple, Microsoft a récemment acquis une start-up Israélienne spécialisée dans la détection de motifs suspects dans les flux de paquets qui circulent sur nos réseaux, en utilisant des algorithmes de *machine learning*.

L'apprentissage bayésien, que nous avons cité plus haut, est également utilisé dans de multiples applications, dans la finance, dans le diagnostic médical, mais aussi dans les réseaux pour garantir des comportements adaptatifs. Un exemple célèbre est celui des « turbo codes » utilisés dans les réseaux cellulaires 3G et 4G. Le filtrage antispam, que nous utilisons tous les jours, est une application classique de l'apprentissage bayésien.

l'IA et le Web

Le développement du Web est lié de façon intime au développement de l'intelligence artificielle et de l'apprentissage. Les moteurs de recherche sont des exemples d'algorithmes « intelligents » qui utilisent des approches de plus en plus riches et de plus en plus de techniques sémantiques issues du traitement du langage naturel. Une des plus grosses applications du *machine learning*, qui nous touche tous les jours, est son utilisation pour le choix et le placement des publicités sur le Web. C'est historiquement, avant le développement du *Deep Learning*, une des premières applications à avoir bénéficié de l'explosion de la puissance de calcul dans un mode distribué. Deux exemples célèbres, que nous avons abordés dans le rapport sur le *Big Data* [1] sont la plate-forme *AdSense* de Google, qui choisit les publicités associées aux mots-clés de nos recherches, et la plate-forme de *retargeting* de Criteo.

L'autre domaine historique du Web qui est une application directe des méthodes de l'intelligence artificielle et de l'apprentissage est la recommandation. Il y a

bien sur les exemples phares comme la recommandation de produits d'Amazon ou de Netflix. Dans les deux cas, les algorithmes ont considérablement évolué en dix ans pour devenir de plus en plus pertinents et constituer une compétence différenciante et stratégique de ces entreprises. Mais il existe des moteurs de recommandation dans la majorité des sites et applications que nous utilisons tous les jours. Qu'il s'agisse de contenus, de services ou de produits, de mise en relation avec des personnes ou des fournisseurs, les algorithmes de recommandation sont omniprésents dans nos vies numériques.

Perception et Deep Learning

Les progrès spectaculaires du *Deep Learning* que nous avons évoqués en introduction ont également déjà trouvé des applications dans notre usage quotidien. La reconnaissance de visage, la reconnaissance de scènes sont progressivement incorporées dans les réseaux sociaux et dans les logiciels qui nous permettent de gérer nos bibliothèques d'images. De façon symétrique, et moins visible pour l'utilisateur, ces méthodes se déploient dans la cybersurveillance, pour transformer des réseaux passifs de surveillance en réseau actifs³⁸. L'arrivée de la conduite autonome, sous la forme progressive des différents niveaux d'autonomie, repose en premier lieu sur la *machine vision* qui est probablement la première application du *Deep Learning*³⁹. Par exemple, la reconnaissance des panneaux de signalisation et des piétons est déjà disponible sur certaines voitures, sans compter l'exemple emblématique de Tesla qui repose également sur les mêmes technologies. La reconnaissance et l'analyse d'image sont appelées à jouer un rôle essentiel dans l'automatisation des processus de production

////////////////////////////////////

- 38 Dans le domaine grand-public, les caméras augmentées par des capacités de reconnaissance intelligente de visages et de situations sont déjà disponibles, comme par exemple la caméra « Welcome » de la société Netatmo.
- 39 Sur ce thème, il faut absolument écouter les conférences de Fei-Fei Li (professeur à Stanford), telle que : https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_pictures?language=fr, ainsi que la leçon inaugurale de Yann Le Cun au Collège de France.

(un des domaines de l'Industrie 4.0⁴⁰). Grâce à la baisse spectaculaire du prix des caméras numériques, il devient possible de numériser les processus de fabrication puis d'entraîner des réseaux neuronaux pour automatiser une partie importante du contrôle qualité. Ces méthodes se combinent très bien avec les drones et les applications de surveillance automatisées se multiplient, par exemple dans l'industrie pétrolière pour la surveillance des pipelines⁴¹.

Au-delà de la vision ou des exemples emblématiques comme le jeu de Go⁴², le *Deep Learning* a des multiples applications comme outil de classification et de reconnaissance de formes. Lors d'une session du programme exécutif de la *Singularity University*, la majorité des fonds d'investissement ont déclaré utiliser l'apprentissage et le *Deep Learning* pour optimiser la gestion de leurs actifs. Comme cela a été évoqué précédemment, le *Deep Learning* est devenu l'outil de choix pour le traitement de langage naturel, en particulier la traduction. Sans surprise, il a rapidement rejoint la boîte à outils des applications de détection de fraude, dans tous les domaines. Autrement dit, avec la généralisation de l'accès aux outils de *Deep Learning* en *open source*, l'ensemble des applications évoquées dans les sections précédentes vont pouvoir bénéficier de la puissance des réseaux neuronaux profonds comme outil d'analyse, du moment que des volumes suffisants de données traitées sont disponibles.

Même si le lien entre les neurosciences et les réseaux neuronaux est ténu, les progrès constants dans la compréhension du cerveau sont une source constante

- ////////////////////////////////////
- 40 Sur ce sujet, lire l'excellent article de Forbes sur la stratégie de GE, *How AI And Machine Learning Are Helping Drive The GE Digital Transformation*: <https://www.forbes.com/sites/ciocentral/2017/06/07/how-ai-and-machine-learning-are-helping-drive-the-ge-digital-transformation/#30a4cc01686b>. Il est particulièrement intéressant de noter la stratégie de plate-forme d'intégration de données, nécessaire pour intégrer les données en provenance de ses partenaires.
- 41 De façon plus générale le « machine vision » est un ingrédient fondamental de la réalité augmentée. Lire par exemple *Why Augmented Reality Is About To Take Over Your World* : https://www.buzzfeed.com/alexkantrowitz/why-augmented-reality-is-about-to-take-over-your-world?utm_term=.bwl4vglybp#.kp2KRmSjWM.
- 42 La victoire d'*AlphaGo* a été une surprise, qui nous éclaire autant sur le Go que sur l'intelligence artificielle (sur la nature de complexité du jeu).

d'inspiration [24] — les réseaux convolutifs en sont un exemple. Demis Hassabis, le CEO de Deep Mind, considère que l'Intelligence artificielle doit encore beaucoup apprendre de l'apprentissage humain pour progresser⁴³.

Traitement des connaissances

Les applications de l'Intelligence artificielle sont également très nombreuses dans le domaine de la représentation et du traitement des connaissances. L'application la plus courante est bien sûr l'utilisation des moteurs de recherche que nous avons déjà évoquée — depuis ses déclinaisons grand public tel que *Google Search* jusqu'aux applications spécialisées dans les bases de connaissances d'entreprises. La plate-forme *IBM Watson* est emblématique de ce type d'usage. On trouve des applications de *Watson* dans tous les types d'industries, qui permettent à des clients ou à des employés d'avoir accès plus simplement et plus facilement à des bases de connaissances, qu'il s'agisse du domaine de la santé, de la finance, des médias ou du voyage.

L'utilisation de la plate-forme *Watson* dans la santé recouvre différents types d'usages, correspondant à la nature hybride de cette plate-forme. Cela va de la reconnaissance de formes dans les images pour assister au diagnostic médical — par exemple la détection et la qualification de tumeurs cancéreuses — à la reconnaissance de « motifs » en épidémiologie pour analyses des bases de données de tests de médicaments en passant par l'utilisation des capacités de traitement de langage naturel pour assister les praticiens dans leur recherche d'information. Plus généralement, la médecine est un domaine historique de l'utilisation de l'intelligence artificielle, puisque les premiers systèmes experts ont été écrits pour assister le diagnostic médical⁴⁴.

43 Lire l'article du MIT Review *Google's AI Guru Says That Great Artificial Intelligence Must Build on Neuroscience* : <https://www.technologyreview.com/s/608317/googles-ai-guru-says-that-great-artificial-intelligence-must-build-on-neuroscience/>

44 Pour aller plus loin sur ce thème « IA et médecine », lire par exemple l'article de Forbes *Artificial Intelligence in medicine – the rise of the machines* : <https://www.forbes.com/sites/paulhsieh/2017/04/30/ai-in-medicine-rise-of-the-machines/#54dd894dabb0>

Le traitement intelligent du langage et des connaissances trouve également son application dans le domaine du droit. Un exemple est fourni par Ross, *The super intelligent attorney*, un logiciel créé par une start-up canadienne qui utilise la plate-forme IBM Watson⁴⁵. L'avis des spécialistes est que nous sommes encore très loin d'un avocat autonome, mais que Ross est remarquablement efficace dans la phase de « *discovery* ». Un autre exemple intéressant de l'application directe des techniques de TLN au droit est la solution *Rocker Lawyer*, un système intelligent de production de documents légaux⁴⁶ qui a remporté un vif succès comme en témoigne sa croissance. L'Intelligence artificielle a encore beaucoup de progrès à faire pour pouvoir permettre une complète automatisation des procédures juridiques, mais son utilisation en tant qu'outil est déjà une réalité⁴⁷.

Il existe de multiples autres domaines d'application

Il existe de multiples exemples d'application de l'intelligence artificielle dans nos vies quotidiennes, et ce rapport n'a pas la prétention de les énumérer. Yves Demazeau, de l'AFIA, nous a cité des exemples dans la maison intelligente, des compagnons artificiels d'assistance de vie, dans la gestion des réseaux, dans les jeux, les réseaux sociaux et l'e-commerce. Parmi nos experts, le témoignage de François Laburthe de Amadeus nous a permis de mieux comprendre comment l'arrivée de l'IA et de l'apprentissage transforme le modèle métier. Amadeus développe et exploite une plate-forme biface dans le marché de la réservation de voyages aériens (3 milliards de passagers-vols par an). L'Intelligence artificielle permet d'augmenter la valeur grâce à une segmentation beaucoup plus fine d'un marché très complexe et très volatil (changements constants des prix et des attentes). Les méthodes utilisées sont principalement des méthodes d'apprentissage, dans un cycle constant d'optimisation et au sein d'une pratique d'intégration et d'ingénierie des données, un sujet sur lequel nous reviendrons

45 Voir *Meet Ross, the Watson powered super intelligent attorney* : <http://www.itbusiness.ca/news/meet-ross-the-watson-powered-super-intelligent-attorney/53376>

46 Voir leur site web <https://www.rocketlawyer.com/> ou la page Wikipédia qui leur est consacrée.

47 Lire par exemple *A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet* : <https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html?mcubz=1>

dans la prochaine partie. Benoît Rottembourg d'Eurodécision nous a confirmé l'arrivée de ces méthodes d'Intelligence artificielle et apprentissage dans le *pricing* et le *revenue management*.

L'intelligence artificielle trouve également de multiples applications dans l'amélioration des interfaces et du dialogue homme-machine. Plusieurs de nos experts ont souligné que les technologies de reconnaissance de la parole, des images et du langage naturel sont parfaitement indiquées pour faciliter le dialogue avec la machine et rendre l'interaction plus simple pour des nouvelles catégories d'utilisateurs. Au-delà du mode de communication, donner une capacité de « raisonnement de sens commun » à nos applications est la meilleure façon de simplifier cette interaction et de supprimer des « irritants » qui sont encore courants dans notre dialogue homme-machine. Les algorithmes d'apprentissage rendent les interfaces plus adaptatives : comme l'a souligné Norm Judah de Microsoft, « plus on utilise Cortana — l'assistant virtuel de Microsoft — plus cela marche ! ». Une des promesses de SNIPS, une des start-up françaises spécialisée dans l'utilisation de l'Intelligence artificielle et de l'apprentissage, est que « l'IA peut faire disparaître la technologie de l'expérience ».

Pour conclure ce tour d'horizon incomplet, il faut souligner la combinaison de l'intelligence artificielle et de « l'économie des intentions »⁴⁸. La reconnaissance des intentions (comprendre ce que le consommateur veut ou cherche, sans qu'il doive l'exprimer) est un des sujets « chauds » du moment. Les entreprises redéployent leurs capacités analytiques pour la découverte des « *insights* », en utilisant des méthodes d'intelligence artificielle. SNIPS, ou Neura — une start-up Israélienne dans laquelle le groupe AXA a investi — proposent de collecter et croiser des événements clients de bas niveaux, produits par exemple par le smartphone, des objets connectés ou la navigation Web, pour reconstruire un contexte puis une intention. Ce qui rend ce domaine passionnant c'est qu'il com-

48 Sur ce sujet, deux des livres références sont *The Intension Economy* de Doc Searl [5] et *Le Marketing Synchronisé* de Marco Tinelli [18].

bine toutes les technologies que de nous venons de présenter, du *Deep Learning* appliqué aux données collectées par des capteurs du téléphone, au traitement de langage naturel (un des sous-thèmes les plus « chauds du moment » étant la capacité d'analyser des textes courts — comme des échanges avec un « chatbot » — pour détecter et reconnaître l'intention du client).

INTELLIGENCE ARTIFICIELLE, APPRENTISSAGE ET MÉGADONNÉES

Il existe une véritable codépendance entre IA et *Big Data* : l'IA a le plus souvent besoin de très grosses masses de données et la masse de données exige des nouvelles techniques pour les traiter. La plupart des méthodes modernes s'appuient sur l'apprentissage et sur l'analyse d'un volume important de données. C'est absolument crucial pour le *Deep Learning*, cela a été souligné par l'ensemble des experts que nous avons consultés⁴⁹. Ray Kurzweil explique qu'*AlphaGo* a d'abord été entraîné sur des millions de parties et a atteint un bon niveau national, mais qu'il a fallu l'analyse de milliards de parties (contre lui-même, en utilisant l'apprentissage par renforcement) pour atteindre le meilleur niveau mondial. De façon symétrique, l'avalanche des données disponibles fait qu'il est nécessaire de trouver des méthodes automatiques et intelligentes pour les traiter [1]. C'est une autre raison qui pousse les chercheurs à s'intéresser à l'apprentissage non-supervisé et à l'introduction de formes de « créativité » dans les méthodes d'intelligence artificielle.

Même s'il existe des frontières et des intersections multiples, l'intelligence artificielle, l'apprentissage et le traitement des données sont des disciplines différentes. Le *machine learning* n'est pas une forme d'intelligence artificielle, c'est une approche générique avec de multiples applications. Le domaine de

49 Dans certains cas, lorsque le problème est très contraint (exemple d'un jeu), on peut utiliser des techniques de génération de données pour produire les corpus qui sont ensuite traités avec des réseaux neuronaux profonds.

l'analyse des données est également distinct, mais avec une double relation puisque nous venons de rappeler que les méthodes qui ont produit les meilleurs résultats depuis 2010 sont des méthodes fondées sur l'utilisation massive de données. La figure 5 explique ce positionnement *de façon simplifiée*. La relation entre l'IA et les données s'inscrit dans un mouvement plus général, qui consiste à « programmer à partir des données » [cf. *Data as code* [1]] en faisant « émerger » le code des algorithmes à partir d'apprentissage. Dans cette perspective, l'approche des réseaux neuronaux est vue comme une nouvelle technique de programmation, très générale, adaptée à des problèmes pour lesquels l'approche traditionnelle de conception des algorithmes est peu satisfaisante. Notons que les algorithmes de *classification*, que l'on peut considérer comme des algorithmes d'apprentissage non supervisés, existent sous de nombreuses formes, des approches purement statistiques comme le *clustering*, aux méthodes de représentation de connaissances.

En revanche, une vision à plat des disciplines cache l'hybridation riche qui rend les frontières très difficiles à distinguer. La figure citée donne une fausse impression de simplicité et de classification, tandis que la réalité est plus complexe. Avant même l'apparition du *Deep Learning*, il existait une hybridation profonde entre les algorithmes de recherche opérationnelle, les méthodes de l'intelligence artificielle et les techniques d'apprentissage. Ce type d'approche hybride est particulièrement nécessaire pour produire des systèmes intelligents autonomes, tout comme pour construire des « résolveurs de problèmes » plus généraux que la plupart des systèmes étroits (cf. le *Todai robot* cité plus haut). Ces systèmes utilisent des méta-heuristiques que l'on peut qualifier de « curiosité », d'« émotions » et de « conscience » même s'il s'agit d'une reconstitution / qualification anthropomorphe plutôt que d'une forme d'imitation de l'homme. La « curiosité » correspond à la programmation de capacités « génératives » utilisant des techniques d'exploration aléatoires (1.2.3). Comme le remarque Michio Kaku [20], cette « curiosité » est une excellente méta-heu-

ristique, elle permet d'améliorer la capacité des systèmes « intelligents » à mieux prévoir en valorisant la « surprise » qui n'est autre que la différence entre la réalité et la prévision. Les « émotions » sont un système de valorisation de l'apprentissage qui peut être piloté par une approche évolutionnaire (ce qui est fait par la nature dans l'évolution des espèces) ou toute autre forme de méta-heuristique dans une double boucle de méta-apprentissage⁵⁰. Le concept de conscience artificielle a fait couler beaucoup d'encre dans le contexte des robots intelligents et autonomes. Sous sa forme la plus simple, il ne s'agit que de doter un système intelligent de la capacité de se percevoir dans son environnement, de percevoir le résultat de ses actions et de pouvoir les anticiper. Ce type de capacité est en fait classique dans les approches classiques de robotiques et d'IA. La révolution commence par les vraies capacités de perception dont nous pouvons maintenant doter les systèmes autonomes grâce au *Deep Learning*.

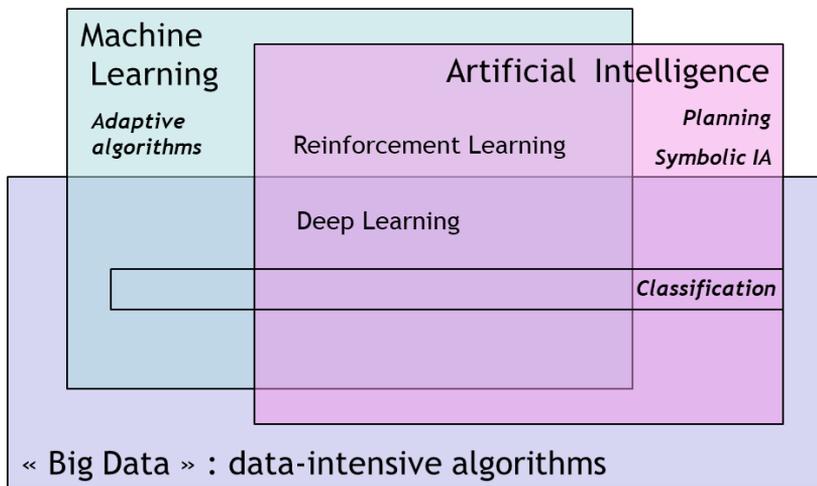


Figure 5 : Liens forts entre IA, Big Data et Apprentissage

50 Ici aussi, pour aller plus loin, il faut lire le livre de Michio Kaku [20] ou à défaut l'introduction simplifiée dans le billet *Event-driven architecture and biomimicry* : <http://informationssystemsbiology.blogspot.fr/2015/12/event-driven-architecture-and-biomimicry.htm>

Le débat sur ce que serait une conscience artificielle gagne en clarté dès que l'on accepte une définition progressive et graduelle, au lieu d'une vision binaire.

Il est plus simple de penser la « conscience artificielle » d'un système intelligent comme une capacité de perception et de raisonnement sur ce système et ses actions. La voiture autonome qui conduit sur la base d'un réseau neuronal couplé à ses caméras n'a pas forcément de « conscience » d'elle-même ; mais si elle commence à anticiper les conséquences de son déplacement en fonction de sa masse et des objets avec lesquels elle pourrait entrer en collision, c'est un premier degré de conscience. Cette approche par degrés s'applique à de multiples définitions possibles de la conscience. Par exemple [19], la conscience peut être vue comme la fondation de la subjectivité [ce qui permet de dire « je »]. Un système intelligent capable de distinguer différentes opinions et points de vue (ce qui est un des objectifs des logiques modales qui sont développées en Intelligence artificielle depuis des décennies) sera facilement capable d'identifier son propre point de vue. Un robot autonome qui fonctionne dans une boucle d'apprentissage non supervisé va construire une représentation de son environnement et de ses propres actions et objectifs. Dans tous ces cas, il ne s'agit pas de conscience au sens existentiel du terme, mais d'une capacité qui — d'une façon semblable au test de Turing — va de plus en plus « ressembler » à la manifestation d'une conscience, au fur et à mesure de la progression des capacités de perception et de raisonnement.

POURQUOI UN RAPPORT DE L'ACADÉMIE DES TECHNOLOGIES ?

Les rapports récents

Il existe une profusion de rapports sur l'IA disponible en ligne, nous allons ici aborder brièvement quelques rapports récents qui incluent des recommanda-

tions⁵¹. Nous avons retenu les cinq rapports récents qui nous semblaient les plus importants :

- le rapport d'INRIA « *Artificial Intelligence — Current Challenges and Inria's engagement* » [13] est un excellent rapport de vulgarisation et une bonne introduction à l'Intelligence artificielle et ses applications. Il permet d'appréhender à la fois la richesse du domaine de recherche et certains des axes prioritaires du moment et des années à venir. Parmi ceux-ci, on peut souligner : la coopération IA-Humain (*human in the loop*), l'apprentissage non-supervisé, la validation et certification, l'analyse des vidéos ;
- le rapport de la *Maison-Blanche* « *Preparing for the future of Artificial Intelligence* » [29] paru en octobre 2016 mérite assurément d'être lu pour comprendre le point de vue américain, qui est à la fois volontariste et optimiste — il s'intéresse dès son début aux applications de l'IA pour le bien commun. Il contient des déclarations fortes, telles que « *Artificial Intelligence has the potential to help address some of the biggest challenges that the society faces* ». La présentation générale de l'Intelligence artificielle est en ligne avec les premières pages de ce document. Les recommandations sont avant tout des recommandations thématiques sur ce que l'IA peut faire pour le gouvernement, pour la société civile, pour le développement de l'économie et de la compétitivité. Il s'agit d'une priorité fondamentale pour le gouvernement américain qui la décline sur le long terme « *AI holds the potential to be a major driver of economic growth and social progress* », par exemple en matière de recherche « *(we) should prioritize basic and long-term AI research* » ou de politique d'ouverture des données publiques. Cette vision résolument optimiste n'est pas sans conscience des risques et des besoins de développer une gouvernance traitant les problèmes de sécurité et de contrôle ;

//////////
 51 Parmi les très bons documents qui ne rentrent pas dans cette catégorie, il est recommandé de lire : *One Hundred Year Study on Artificial Intelligence*, Stanford University, 2014. Voir aussi l'article *Artificial Intelligence and Life in 2030* pour une bonne prospective sur les applications de l'IA.

- le rapport le plus important en France est le Rapport de synthèse « France IA » [10], paru en mars 2017. Il commence par un excellent panorama de la recherche scientifique en IA, qui approfondit chacun des domaines ou concepts que nous avons exposé trop brièvement dans le présent document. Il continue par une prospective thématique des applications de l'IA, dans de nombreux domaines. Le seul défaut de ce rapport, fruit d'un considérable travail collectif, est d'être quelque peu « touffu »... En contrepartie il contient de nombreuses « pépites » comme la très bonne section sur l'IA et le droit. Nous allons revenir sur les recommandations proposées en synthèse, elles forment un consensus à la grande majorité des rapports. Il faut souligner d'une part le ton très positif du rapport sur l'Intelligence artificielle, très similaire au rapport de la Maison-Blanche et l'optimisme du rapport sur la situation concurrentielle de la France, sur laquelle nous reviendrons en conclusion;
- l'OPECST a également publié en 2017 un rapport intitulé *Pour une intelligence artificielle utile, maîtrisée et partagée* [22]. Ce rapport contient également de bonnes recommandations en faveur de l'Intelligence artificielle : « ne pas céder à la tentation de réguler trop vite », « favoriser des algorithmes sûrs, transparents et justes », mais qui peuvent sembler rester un peu trop générales. Néanmoins, ce rapport contient beaucoup de recommandations très justes sur l'importance de la formation, qui est un autre point qui fait consensus chez les experts. Un des objectifs du rapport — visible dans son titre — est de trouver un équilibre entre un plaidoyer favorable à l'IA et la prise en compte des risques et des inquiétudes sociétales. En conséquence, une partie importante du rapport traite de ce que l'IA saura ou ne saura pas faire, quelquefois avec un peu trop de certitudes. Nous allons y revenir dans les parties suivantes : il est difficile — et souvent faux — de dire ce qui sera possible et quand⁵²;

////////////////////
52 Ce point fondamental est très bien expliqué dans le livre de Max Tegmark [19] qui est une des meilleures références sur le sujet. À l'inverse, nombre de livres récents qui sont parus pour dénoncer les excès du discours sur l'intelligence artificielle souffre du même travers : celui de penser que nous sommes capables aujourd'hui de dire avec autorité ce qui sera ou ne sera pas possible en 2050.

- pour finir, il nous semble important de signaler l'excellent rapport du CGEJET — *Modalité de régulation des algorithmes de traitement des contenus* [14]. Même si ce rapport est sur un sujet plus étroit (lié au domaine du contenu sur le Web) qui n'est pas à proprement parler le sujet de l'intelligence artificielle, il pose avec beaucoup d'acuité et de pertinence la question du contrôle des « algorithmes intelligents » et constitue la meilleure référence sur ce sujet. Ce rapport se place en effet d'emblée dans le contexte des nouvelles technologies logicielles : algorithmes construits par émergence à partir des données, besoin de raisonner en fonction des écosystèmes d'acteurs, importance de l'intention et capacité à l'auditer, nécessité sociétale des explications. Les auteurs soulignent des points très intéressants, comme la capacité à propager des biais à des algorithmes au travers des jeux de données d'apprentissage. Les recommandations⁵³ de ce rapport portent sur la création de plates-formes de tests, sur la certification et la communication autour des algorithmes et sur le besoin de formation des opérateurs des algorithmes.

Ce qui fait consensus sur le domaine de l'Intelligence artificielle et de l'apprentissage

Ce qui suit est une liste des points qui font consensus dans l'ensemble des rapports que nous venons d'évoquer :

- Il s'est passé quelque chose depuis 2010, il y a une accélération spectaculaire des résultats en termes d'intelligence artificielle qui est plus proche de la révolution que de l'évolution — tirée en premier lieu par une révolution technologique qui a entraîné une évolution des algorithmes. Il y a un consensus pour dire que cette inflexion est liée à la combinaison de l'augmentation très brutale des performances (la combinaison de l'accélération exponentielle

53 Ces recommandations sont à rapprocher des sept principes pour la transparence algorithmique et la responsabilisation de l'ACM : https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

de la Loi de Moore et de l'utilisation de processeurs plus efficaces comme les GPU), la constitution de très grands corpus de données (et la capacité à les stocker et les gérer, ce qui lie l'IA et le *Big Data*) et enfin les progrès algorithmiques, en particulier sur les algorithmes de *Deep Learning*. Cette rupture a provoqué un afflux massif des investissements, tant dans le monde du capital-risque que chez les grands acteurs numériques. Tout le monde s'accorde à dire que ces investissements sont très significatifs, parce que l'Intelligence artificielle est également une pratique, ce qui rend son développement fortement corrélé aux moyens disponibles. Les investissements massifs (en milliards de dollars) des GAFAM — et IBM dans une moindre mesure - et équivalents Chinois (Baidu, Alibaba, Tencent) interrogent l'Europe sur notre capacité à tenir notre place dans cette course.

- Il existe un vrai enjeu de compétitivité pour nos économies et entreprises. Le rapport de la Maison-Blanche est très explicite : « *Advances in AI hold incredible potential to help United States to stay on the cutting edge of innovation* ». Il faut penser à l'intelligence artificielle comme à une boîte à outils, un *enabler* de toutes les autres formes d'activité et non pas comme à une fin en soi. D'un point de vue économique, le sujet de l'imitation de l'homme est l'arbre qui cache la forêt du « *general problem solver* ».
- Le développement de l'IA passe par la mise à disposition de gros volumes de données. Les GAFAM sont précisément les détentrices de données personnelles et qualifiées de plusieurs centaines de millions d'utilisateurs (plus d'un milliard pour Facebook). Une politique de soutien du développement doit favoriser la constitution de jeux de données pour l'entraînement des futures IA, et proposer une infrastructure qui mette ces données à la disposition du plus grand nombre d'acteurs innovants.
- Il y a un fort besoin de formation autour de l'Intelligence artificielle, pour tous, depuis les étudiants jusqu'aux citoyens en passant par les managers. Comme le propose le rapport de la Maison-Blanche : « *Educate and train Americans for the jobs of the future* ». Ce besoin de formation est lié à la vision *enabler*

de l'IA : il s'agit de permettre à l'ensemble de nos économies — des producteurs aux consommateurs — de tirer le meilleur parti de cette opportunité, pour améliorer l'ensemble de nos produits et services⁵⁴.

- Favoriser le développement de l'Intelligence artificielle consiste à favoriser l'émergence d'un écosystème d'acteurs. Il existe une chaîne de valeur complète et complexe de fournisseurs depuis le développement des briques techniques jusqu'à l'expérience fournie au consommateur. Dans cet écosystème coexistent des grandes entreprises qui bénéficient des effets d'échelle et des petits acteurs plus agiles. La notion d'écosystème s'étend au logiciel — en particulier autour de l'open source et du monde des API qui offrent des fonctions de l'IA « *as a service* » — dans une approche ouverte, telle IBM avec *Watson*, mais aussi dans une approche de plateforme interne comme Salesforce avec *Einstein*. Le concept d'écosystème s'applique parfaitement à l'univers des données, avec des exemples ouverts comme les stratégies *Open data* des gouvernements ou institutions, mais également des stratégies audacieuses de plates-formes « ouvertes » d'intégration de données comme celle de GE⁵⁵. Construire, favoriser et animer un écosystème⁵⁶ reste un travail difficile et de longue haleine : il faut favoriser les contacts entre académiques, start-up, entreprises, fournisseurs, administration... et créer des lieux et des événements — tels que les hackathons — qui favorisent des échanges informels.
- On trouve dans l'ensemble des rapports français la recommandation de favoriser le transfert depuis le monde de la recherche et l'innovation vers les

54 Le rapport qui vient de sortir au Royaume-Uni *Growing Artificial Intelligence in the UK*, de Wendy Hall et Jérôme Pesenti, comporte des recommandations très précises en matière de formation qui pourraient être avantageusement déclinées en France. <https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk>.

55 Lire l'article *How AI And Machine Learning Are Helping Drive The GE Digital Transformation* : <https://www.forbes.com/sites/ciocentral/2017/06/07/how-ai-and-machine-learning-are-helping-drive-the-ge-digital-transformation/#222ef5e71686>

56 Sur ce concept, l'article de référence reste celui de Michel Ferrary et Marc Granovetter sur la silicon valley : *The Role of Venture Capital Firms in Silicon Valley's Complex Innovation Network*.

entreprises. Il serait difficile d'être en désaccord avec un tel objectif, mais cela reste souvent à l'état de déclaration incantatoire. Très modestement, les éléments de la suite de ce rapport se veulent des contributions à un tel objectif.

Conformément à ce qui a été dit en introduction, ce présent rapport va aborder ces points différents dans les trois parties suivantes.

Quelle Intelligence artificielle pour quels problèmes ?

COMPRENDRE LA BOÎTE À OUTILS

Le premier objectif de ce rapport est de contribuer à démystifier l'IA pour permettre au lecteur de mieux comprendre ce qui est possible et ce qui ne l'est pas encore. La plupart des experts que nous avons interrogés ont insisté sur le fait qu'il y avait une double confusion sur ce qu'était l'IA et sur ce qui était possible aujourd'hui. La boîte à outils est plus large que les entreprises ne le comprennent, et ce phénomène est accentué par la presse qui ne s'intéresse — logiquement — qu'à ce qui est perçu comme nouveau. Cette obsession autour du *Deep Learning* a le double inconvénient de faire oublier d'autres techniques dont la mise en œuvre serait plus simple et de faire croire qu'il s'agirait d'une panacée.

Les entreprises qui souhaitent expérimenter des approches d'IA sur leurs problèmes disposent d'une large boîte à outils. Sans surprise, on retrouve les familles de méthodes que nous avons vues dans la section précédente :

- les outils classiques de « *data science* » / « *data mining* ». Ces outils font partie de bibliothèques classiques d'algorithmes, qu'il s'agisse de packages commerciaux ou *open source*. Le livre *Data Science : Fondamentaux et études de cas* [?], qui s'adresse à des étudiants, donne une très bonne

idée des principales méthodes qui sont utilisées par les entreprises du domaine. On retrouve une boîte à outils semblable chez les experts — par exemple la start-up Da Vinci Labs précédemment mentionnée — ou dans les plates-formes adaptatives comme Einstein ;

- les moteurs de règles sont des outils logiciels matures, disponibles depuis de longues années, qui ont fait leur preuve, y compris en combinaison avec d'autres outils comme l'orchestration ou le *Complex event processing*. Le monde de l'IoT (*Internet of Things*) et du big data des senseurs connectés s'appuie sur des architectures « orientée évènement » qui se marient harmonieusement avec des moteurs de règles pour construire des systèmes intelligents « temps réel » ;
- l'automatisation d'agent dans des processus est une variante des moteurs de règles qui inclut des capacités de *scripting* et de traitement de langage naturel pour simuler des agents de façon « intelligente ». Connue sous le nom de RPA (*Robotic Process Automation*), cette technologie est employée dès maintenant par de nombreuses entreprises ;
- les agents conversationnels (*chatbots*) permettent de conduire une interaction avec un client ou un employé en langage naturel, qu'il s'agisse de texte (messages) ou de voix. Ces plates-formes restent simples en termes de capacité de raisonnement et de compréhension profonde des intentions du client. Dans le meilleur des cas, elles permettent de créer des assistants « intelligents » de type Siri ou Alexa, mais cela demande un travail considérable dans le domaine considéré — un très gros apprentissage sur un corpus important d'interactions. Le plus souvent, les *chatbots* se limitent à des dialogues simples sur un domaine étroit, mais cette forme d'interaction apporte déjà une véritable satisfaction des clients dans des cas précis⁵⁷ ;
- Il existe plusieurs plates-formes pour construire des traitements intelligents du langage naturel, depuis l'extraction des émotions jusqu'à l'analyse

//////////
 57 Les chatbots sont pertinents avec des clients qui sont mal à l'aise avec les interfaces graphiques habituelles, ou ceux qui ne sont pas en condition de les utiliser (mains encombrées, mobilité, etc.).

sémantique. La plate-forme IBM est accessible via des API simples, mais de nombreux autres outils, en particulier des outils *open source* plus spécialisés, peuvent être intégrés sans efforts particuliers ;

- les méthodes modernes d'apprentissage par réseaux neuronaux sont facilement accessibles puisque la plupart des algorithmes sont disponibles en libre-service (*Software as a Service*) ou en librairie open source ;
- et pour conclure, il ne faut pas oublier les bibliothèques conséquentes d'outils et de services d'aide à la décision dans les domaines connexes de l'Intelligence artificielle.

L'accès aux techniques d'Intelligence artificielle est fortement facilité par la disponibilité de l'ensemble des « briques de base » en open source. Si ce rapport fait souvent référence à l'open source, c'est qu'il s'agit d'une tendance de fond qu'ont souligné la plupart de nos experts. Dès le départ les domaines scientifiques de l'IA en général et surtout du *data mining* ont donné lieu à de multiples outils en open source. Cette tendance s'est accentuée depuis quatre ans, puisque tous les grands acteurs (les GAFAM) ont décidé de rendre accessible, à tour de rôle, une partie importante de leur patrimoine logiciel en *machine learning* — citons par exemple : *Torch* chez Facebook (et également *Big Sur*, pour l'utilisation des GPU), *TensorFlow* chez Google, *Cortana* chez Microsoft, *Watson* (en partie) chez IBM et plus récemment une partie des algorithmes d'Apple qui sont disponibles sous iOS. Ceci n'est pas une coïncidence : il y a une conviction partagée que nous n'en sommes qu'au début de l'aventure de l'intelligence artificielle et que l'écosystème gagnant sera celui qui attirera le nombre le plus large de talents et de données validées. La conséquence évidente pour les entreprises est qu'il faut en profiter : il s'agit d'assembler des briques qui sont mises au point en continu par des communautés très larges.

SAVOIR APPLIQUER DES MÉTHODES SIMPLES À BON ESCIENT

Même s'il n'existe pas de règle simple pour savoir quelle méthode appliquer, un peu de bon sens permet de comprendre quelles approches conviennent mieux à chaque situation, en fonction de la question que l'on se pose et des données disponibles. La figure 6 propose une vision très simplifiée — donc réductrice — de l'influence de deux critères majeurs sur le choix de l'approche. Nous avons déjà évoqué l'importance des données pour les méthodes fondées sur l'apprentissage, et en particulier pour le *Deep Learning*. L'autre axe proposé est la nature de la question ou du problème à résoudre. Les questions ouvertes sont mieux traitées par l'utilisation d'agents intelligents, sous forme de simulation exploratrice, où par l'utilisation de techniques sémantiques couplées à de très grandes bases de connaissances lorsqu'il s'agit de répondre à une question posée. Comme rappelé dans la section précédente, les méthodes statistiques de classification et prédiction ont une large applicabilité. Notons également que la Loi de Moore a apporté un bénéfice substantiel pour la plupart de ces méthodes, qui ont fait des progrès importants sur la dernière décennie.

Il faut insister sur le fait que les méthodes simples fonctionnent bien pour une grande classe de problèmes, et que l'IA ne doit pas être réduite au *Deep Learning*. Lors des entretiens autour des méthodes de *Big Data* [1], les experts de Critéo, Facebook et Google ont insisté sur le fait qu'ils utilisaient des algorithmes simples, mais avec un très grand nombre de paramètres, sur de très grands volumes de données et au moyen d'un processus complexe d'apprentissage et de mise au point — là où réside le savoir-faire acquis par l'expérience. Nous avons illustré le *Deep Learning* par le cas particulier des *Convolutated neural networks* (CNN) car ils jouent un rôle essentiel pour les problèmes de reconnaissance d'image — et pour les problèmes de *Machine vision* ; il n'y a aucun doute que les CNN sont, dans l'état actuel de nos connaissances, LA méthode à appliquer. Terminons par une citation de Neil Jacobstein : « *On most problem small data is enough, but deep learning craves for data* ».

L'utilisation de l'Intelligence artificielle pour prédire demande un peu de connaissances en statistiques et surtout de la rigueur dans les protocoles de test. Cet énoncé devrait être d'une grande banalité, pourtant l'engouement pour l'IA et l'apprentissage conduit des entreprises à tomber dans des pièges classiques, tels que « *the curse of dimensionality* », les « *spurious correlations* » ou le classique « *overfitting* »⁵⁸. Toutes ces manifestations viennent d'une même cause racine : un ensemble de données dont les caractéristiques sont moins nombreuses que les paramètres du modèle d'apprentissage. Les algorithmes vont alors trouver des façons d'adapter « de manière forcée » le modèle, avec une très bonne performance, mais une absence de robustesse et donc aucune valeur de prévision. Lorsqu'il y a un très grand nombre de variables, un grand nombre de corrélations apparaissent de façon fortuite (*spurious*). Plus les algorithmes sont sophistiqués, plus la Loi de Moore met à notre disposition des machines puissantes, plus il est facile de tomber dans ces travers. Il existe de nombreux protocoles de test pour les éviter, mais il faut des compétences statistiques pour comprendre et qualifier les premiers résultats. De nombreux exemples emblématiques soulignent ce problème : la prédiction de « *Google Flu* » [1] ou bien des fausses corrélations entre des pathologies rares et des gènes (trop de gènes, un échantillon petit, les conditions idéales pour des corrélations fortuites) [27]. Pour donner un exemple complètement différent, les méthodes d'apprentissage des réseaux neuronaux sont peu robustes au bruit : les calculs en simple ou double précision ne donnent pas les mêmes résultats et il est souvent nécessaire de faire les calculs en double précision pour obtenir des solutions robustes. Un de nos experts nous a signalé qu'un certain type d'algorithme de propagation de gradient diverge lorsqu'il est porté sur un GPU. De façon générale, l'utilisation de méthodes « randomisées », depuis l'apprentissage jusqu'à la simulation Monte-Carlo ou la théorie des jeux évolutionnaire, exige du recul, de la rigueur et un savoir-faire « scientifique »⁵⁹.

58 La plupart des ouvrages cités dans la bibliographie traitent de ces problèmes, mais peu le font de façon aussi ludique que *Everybody lies : Lies : Big Data, and What the Internet Can Tell Us About Who We Really Are* [27]

59 Sur ce sujet, lire l'article de Martin Goodson *10 ways your data project is going to fail* : <http://>

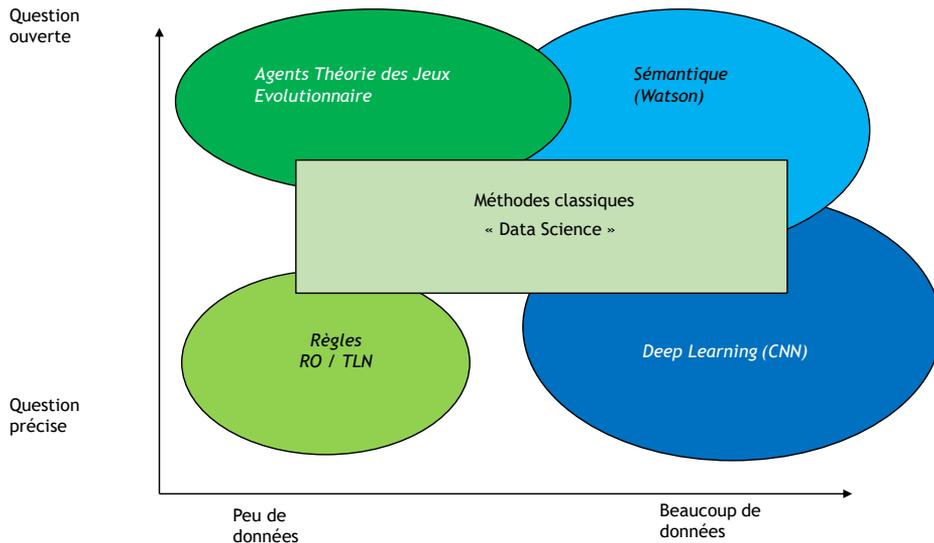


Figure 6 : une vision simplifiée des approches IA en fonction des situations⁶⁰

La mise en œuvre de méthodes d'Intelligence artificielle demande du temps car il y a beaucoup de connaissances pratiques à acquérir. Ce message a été répété par la plupart de nos experts : il y a beaucoup de savoir-faire dans la mise en œuvre — cela dépend néanmoins du type d'approche utilisé. Les ingrédients sont à la disposition de tous, mais la réalisation relève encore de la « cuisine »

www.martingoodson.com/ten-ways-your-data-project-is-going-to-fail/

60 Même si cette figure est simpliste, elle permet par exemple de comprendre qu'un chatbot pour lequel on dispose de peu de données doit être construit en mélangeant règles / scripts et traitement du langage naturel, tandis que si l'on dispose de millions de conversations, une approche par réseaux neuronaux peut donner des résultats beaucoup plus puissants (le chatbot devient un vrai expert).

au sens noble : il faut pratiquer, expérimenter et tester⁶¹. De façon similaire, *Amadeus* nous dit expérimenter des approches sophistiquées en R & D, mais utilise en production des approches plus simples et plus robustes. Le temps est une dimension essentielle du processus : les praticiens adaptent leurs approches au temps dont ils disposent (en général pas plus d'une semaine par itération d'apprentissage) en fonction des machines. C'est pour cela que l'utilisation des GPU est une compétence essentielle à acquérir puisqu'il permet de gagner de deux à trois ordres de grandeurs de temps de calcul.

LE PROCESSUS D'EXTRACTION DE VALEUR DE LA DONNÉE

Mettre en place une démarche d'Intelligence artificielle est un processus qui commence par la collection, la sélection, et la curation de données et qui se termine par une longue validation. Le dessin de la figure suivant illustre les différentes étapes que nous avons évoquées jusqu'ici. La première étape est celle de la curation des données : collecter, enrichir par des métadonnées, en fonction de modèles⁶². Pour de nombreux experts, vu la maturité des méthodes aujourd'hui, une « stratégie IA pour une entreprise » consiste en premier lieu en une « stratégie d'acquisition de données », d'autant plus complexe qu'on ne sait pas aujourd'hui ce que la technologie permettra demain. La deuxième étape consiste à choisir les méthodes que l'on veut tester, dans la boîte à outils du début de ce chapitre. La troisième étape est une étape d'intégration : on retrouve l'importance des compétences informatique et du savoir-faire technologique. Ce dernier est souvent plus important que la compétence scientifique

////////////////////////////////////

- 61 L'exemple fort médiatisé – à juste titre – de la jeune fille indienne qui a réalisé un dispositif médical de diagnostic à partir d'un algorithme d'apprentissage appliqué à une image capturée sur un smartphone illustre bien la diffusion large des techniques. <https://techcrunch.com/2017/08/03/high-schooler-makes-3d-printed-machine-learning-powered-eye-disease-diagnosis-system/>
- 62 La qualité des données, les connaissances associées et le modèle qui permet de les interpréter sont déterminant pour la valeur créée par ce processus. Lire par exemple l'article de Daniel Tunkelang *10 things everyone should know about machine learning* : <https://www.linkedin.com/pulse/10-things-everyone-should-know-machine-learning-daniel-tunkelang/>

pour construire les « meilleurs algorithmes ». La dernière étape est celle de l'optimisation du processus d'apprentissage et de la validation des résultats, dont nous avons souligné l'importance dans la section précédente (il est facile d'être impressionné par les premiers résultats, les « fausses joies » font partie de la vie courante des *data scientists*).

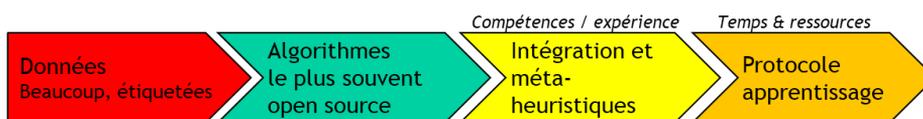


Figure 7 : La mise en place de compétence d'Intelligence artificielle en tant que processus

Le coloriage naïf exprime que les parties les plus techniques et difficiles de ce processus ne sont pas celles du milieu mais bien la collecte et curation des données d'une part, et la validation du protocole d'apprentissage d'autre part.

Ce processus, lorsqu'il est appliqué à un nouveau problème, prend du temps et de la patience car ce temps est indéterminé. Comme nous l'a dit Norm Judah, il faut encore beaucoup de temps pour sélectionner les données et créer un modèle (même si certaines plates-formes comme *Einstein* se proposent précisément d'automatiser cette création de modèle). Un autre expert a souligné que l'apprentissage machine est un processus itératif, qui prend du temps. Il est important de bien distinguer le temps de mise au point de l'apprentissage — celui dont nous parlons ici — et celui de l'exécution de l'algorithme, qui peut être très court et déporté sur des objets ou des machines moins performantes. Cette complexité est plutôt une bonne nouvelle — sous forme de barrière à l'entrée dans de nouveaux métiers — pour les entreprises : elle rend moins crédible le scénario où les compétences d'IA seraient concentrées entre peu d'acteurs (les GAFAM de demain) qui exploiteraient les opportunités — *as a service* — pour l'ensemble du monde. Le paysage des opportunités est extrêmement vaste, et il est beaucoup plus réaliste de voir une fragmentation diversifiée des lieux

d'excellence de l'application de l'IA⁶³. La figure 7 indique un petit cycle itératif de mise au point, mais il en existe un plus large dans le cas de développement de services à partir des données. Dans ce cas, le développement de nouveaux services crée des nouvelles données, et les nouvelles données permettent des nouveaux services. Ceci signifie que le paysage des données se construit également de façon itérative et progressive. Le monde numérique est caractérisé par de nombreux services dont la finalité est la capture de données.

Appliquer une approche d'Intelligence artificielle et d'apprentissage à un processus métier transforme la vision du métier et permet de renouveler les outils et les pratiques. Cette remarque faite par plusieurs des experts sur des domaines différents est très profonde. Elle s'appuie sur plusieurs idées. Premièrement, la vision métier et sa modélisation sont une part intégrante du processus décrit dans la figure précédente⁶⁴. La mise au point itérative des systèmes intelligents alimente la connaissance du métier : « L'utilisation du *machine learning* sur des domaines industriels conduit à approfondir, voire à remettre en cause sa connaissance du métier ». Deuxièmement, l'existence d'une capacité de contrôle intelligente est une grande opportunité de transformation. Par exemple, le métier de plate-forme est profondément transformé par cette intermédiation intelligente. Du point de vue des infrastructures, la capacité de déplacer le contrôle d'anciens systèmes vers des nouveaux pour y intégrer une intelligence artificielle permet de rafraîchir des architectures plus anciennes de systèmes d'information. L'article précédemment cité sur la stratégie de GE, qui utilise ses compétences en apprentissage et une plate-forme ouverte pour

63 Ceci n'empêche pas la création d'écosystèmes de dépendances complexes. Par exemple, dans le domaine du *machine vision*, Google expose les couches basses de ses réseaux de neurones entraînés en reconnaissance d'image, qui peuvent être réutilisées par des équipes qui cherchent à entraîner des réseaux sur des problèmes propres à une industrie, comme des problèmes de maintenance prédictive.

64 Les expériences d'application du machine learning aux domaines industriels montrent qu'il est nécessaire d'ajouter de la connaissance métier aux données brutes collectées par les machines et les senseurs. Pour comprendre l'importance de cette étape de « curation » des données avec des métadonnées issues de la modélisation métier, lire l'article *Applying deep-learning to real world problems* de Rasmus Rothe : <https://medium.com/merantix/applying-deep-learning-to-real-world-problems-ba2d86ac5837>

collecter des données de ses différents partenaires, est une parfaite illustration des deux points précédents : Il existe de multiples opportunités de concentration locale (cf le rapport du CGEJET [14]) pour créer des écosystèmes de données correspondant à un domaine métier, et la maîtrise dans les méthodes évoquées dans ce rapport peuvent se transformer en avantage compétitif très significatif à cause du bénéfice du premier entrant (*first mover advantage*). Il existe donc un risque important de disruption pour les entreprises qui fondent leur différenciation sur l'expertise acquise au travers d'un grand nombre d'expériences. L'utilisation de l'IA peut permettre à des concurrents d'accélérer la vitesse d'acquisition des connaissances (depuis des données de test jusqu'à des verbatims client).

Garder un regard lucide sur les progrès technologiques

UN VRAI RISQUE D'EMBALLEMENT QUI POURRAIT CONDUIRE À UN NOUVEL HIVER

Les publications récentes montrent un emballement sur les possibilités du *Deep Learning* qui est contre-productif. Les articles qui ont accompagné les prouesses citées en introduction ont tendance à exagérer l'applicabilité des réalisations et à oublier la très longue phase d'apprentissage et d'optimisation propre à chacune de ces expérimentations. En conséquence, cela crée un phénomène de « *hype* » et des attentes irréalistes. Cette exagération dans les discours est accentuée par la confusion entre le débat — de nature philosophique — sur les conséquences de l'existence d'une intelligence artificielle générale (AGI) et la réalité des intelligences artificielles étroites (ANI) aujourd'hui. C'est le premier point que nous ont signalé les start-up que nous avons rencontrées : en ce moment, les attentes sur l'IA et l'apprentissage relèvent de la « magie » et la réalité devient vite une déception⁶⁵. Il devient important de réinjecter une

65 Le témoignage d'une jeune chercheuse en machine learning, Chip Huyen, intitulé Confession of a so-called AI expert permet de comprendre cette situation délicate pour les praticiens par

bonne dose de réalisme pour éviter un « nouvel hiver de l'IA ». Les spécialistes de la discipline participent déjà à cette désinflation des discours : par exemple, la conférence *The Future of AI* organisée en 2016 par Facebook et Yann Le Cun avait, entre autres, comme objectif de « refroidir la température », selon l'un de nos experts.

Le domaine des *chatbots* est également atteint par l'inflation des promesses et l'abus du terme « Intelligence artificielle ». Après une phase de surpromesse (*hype*) très forte, il y a un certain désamour autour des *chatbots*, qui ne seraient plus dignes d'être appelés « intelligence artificielle ». L'exemple très médiatique du robot *Tay* de Microsoft, que les internautes ont rapidement détourné puis éduqué à raconter des propos choquants, est cité comme un exemple de l'absence d'« intelligence » de ces *chatbots*. En effet, Ray Kurzweil lui-même, lorsqu'on lui pose la question de l'émergence de *chatbots* réellement intelligents (au sens du test de Turing) donne la date lointaine de 2030 (ce qui peut surprendre puisque ses prédictions sont souvent volontaristes). Le débat n'a pas vraiment de raison d'être : comme la plupart des techniques d'intelligence artificielle, nous n'en sommes qu'au début et les applications ne fonctionnent que dans un cadre étroit, avec une valeur qui est strictement proportionnelle au temps passé à collecter et traiter les données d'apprentissage ou de mise au point (les agents réalisés en quelques semaines ne font guère mieux qu'énoncer des scripts sous forme d'arbre de décision). En revanche, les investissements colossaux réalisés par les GAFAM sur les techniques d'assistants intelligents donnent le double signal qu'il se passe déjà quelque chose et qu'il n'est pas déraisonnable de penser que les plates-formes de *chatbots* vont faire des progrès rapides dans les années à venir.

La complexité du monde rend de nombreux processus imprévisibles, indépendamment de la quantité de données collectée ou de la puissance des algorithmes d'IA qu'on leur applique. Ici aussi, comme dans le chapitre II, il pourrait sembler

rapport à des attentes irréalistes : <https://huyenchip.com/2017/07/28/confession.html>

d'une grande banalité de rappeler que tout n'est pas prévisible⁶⁶. Il n'existe pas de méthode pour prévoir un bruit aléatoire (ce qui est souvent collecté par des senseurs) où un phénomène éminemment complexe comme la météo ou les cours de Bourse (ce qui n'empêche pas bien sûr ni de les étudier ni de chercher à mieux les comprendre). Lorsqu'on se heurte à ce type de difficulté, il est facile de chercher la solution dans la collecte supplémentaire de données, ou l'utilisation d'une plus grande puissance de calcul, si on ne prend pas garde à conserver un jugement statistique prudent sur les résultats. Notre expert sur le *pricing* a souligné la difficulté inhérente de la prévision lorsque les courbes d'élasticité des clients changent constamment (ce que l'on constate) et lorsque les modèles de prévision des comportements sont structurellement assez pauvres⁶⁷. Les clients apprennent — les promotions massives changent les comportements — ce qui produit précisément des phénomènes complexes. Tout ceci souligne une fois de plus l'importance de la phase de validation des résultats et le besoin d'une éducation scientifique dans le domaine statistique.

LE RÔLE CLÉ DE LA PRATIQUE ET DE LA TECHNOLOGIE

L'Intelligence artificielle ne se réduit pas à des données et des algorithmes ; la position concurrentielle de l'Europe est moins favorable que ce qui est souvent dit. L'excellence en IA ne se mesure pas seulement en nombre de chercheurs et de publications. Une partie des évaluations faites au sujet de la France ou de l'Europe ces dernières années est aveuglée par un intérêt exclusif pour la dimension scientifique et algorithmique. Cette remarque est d'autant plus importante

66 La meilleure référence sur ce thème reste le premier livre du médiatique Nassim Taleb *Fooled by Randomness* [21]. Les biais psychologiques qui nous poussent à voir du sens lorsqu'il n'y en a pas (narrative fallacy) sont évidemment renforcés par les approches de *Big Data* et d'apprentissage.

67 On utilise pour cela un indicateur R2 (dit OOS : *out-of-the-sample*) qui mesure la proportion de la variation des observations par rapport à ce qui serait expliqué par le modèle. Les valeurs de 0,6 à 0,7 observées sur des modèles de comportement client se rapprochent du monde de la finance et traduisent la difficulté de la prévision (cf. <http://businessforecastblog.com/out-of-sample-r2-values-for-pvar-models/>)

que **les contributions scientifiques et technologiques françaises en Intelligence artificielle et apprentissage sont très nombreuses** — on trouve massivement des chercheurs et innovateurs français dans les avancées qui sont ensuite souvent attribuées à des sociétés américaines parce que sont celles-ci qui les ont développées à grande l'échelle. Les arguments dans le chapitre II permettent d'apprécier le rôle clé de la pratique, depuis le savoir-faire en termes de collecte et de curation de données, les protocoles de mise au point ou les compétences d'ingénierie logicielle qui sont nécessaires pour manipuler les données et réaliser l'intégration des briques logicielles. **Très souvent, la technologie tire la science** : ce sont les plates-formes technologiques qui permettent de faire des avancées scientifiques et les algorithmes sont issus de l'expérimentation. En conséquence, l'excellence en IA est également une question de moyens et de pratique. Les moyens sont nécessaires pour collecter et assembler les données dans des grands volumes ; ils sont également nécessaires pour fournir aux projets les ressources de calcul et l'expertise en ingénierie. La pratique est — pour employer une métaphore scientifique — l'intégrale des moyens au cours du temps : il faut accumuler des jours de pratique pour développer les différentes expertises que nous avons mentionnées. La création de valeur dans l'Intelligence artificielle ne se fait pas qu'en amont — au niveau de la R & D — elle se fait sur l'ensemble du processus de déploiement — dans l'industrie, les services. La métaphore de l'intégrale souligne l'effet d'accumulation : il n'est pas facile de rattraper un retard en termes de pratiques, il faut faire un effort prolongé sur les moyens⁶⁸.

Le point de départ de toute stratégie en intelligence artificielle est la constitution de larges corpus de données annotées. Ces jeux de données se retrouvent dans l'ensemble des *success stories* de l'IA que nous avons évoquées en introduction, depuis les corpus d'images d'ImageNet jusqu'à la capture massive de trajets par les Google cars. Dans de nombreux domaines, le *crowd-sourcing* est mis en œuvre pour collecter et surtout annoter les informations. Alex Lebrun, le spécialiste

//////////
68 En particulier, et c'est la position de l'Académie des technologies depuis plusieurs années, il faut continuer l'effort d'investissement sur les calculateurs hautes performances (HPC).

mondial des *chatbots* recruté par Facebook, souligne que le développement des *chatbots* « de prochaine génération » sera la conséquence naturelle de la collecte massive des dialogues entre les clients et les services d'assistance d'aujourd'hui. Une partie des problèmes « génériques » vont donner lieu à des « corpus de domaine public », mais pour ce qui est des domaines industriels, le travail de collecte et annotation — en fonction du modèle et de la vision du métier — est un enjeu stratégique⁶⁹.

La position forte de l'Europe sur la protection de la vie privée, en particulier avec la réglementation GDPR, pose une véritable question sur la capacité des entreprises à constituer les corpus de données nécessaires à l'innovation de demain. Laurent Alexandre, membre de l'Académie des technologies, en a fait à juste titre un des points marquants de son intervention à l'université d'été du MEDEF. L'Europe combine une logique — indiscutable — du respect de la vie privée avec une vision déterministe et mécaniste de l'informatique, ce qui la conduit à sanctuariser la « finalité » des actes de collectes de données [1]. Les notions de consentement « explicite » et « positif », ainsi que de « protection des données dès la conception » reposent sur des hypothèses d'une maîtrise complète de la « finalité » des traitements informatiques. Comme nous l'avons vu dans les parties précédentes — et cela est fort bien souligné dans le rapport du CGEJET — l'Intelligence artificielle et l'apprentissage reposent sur des approches émergentes, dans laquelle les algorithmes sont cultivés à partir des données, et pour lesquelles les finalités ne sont validées qu'à la suite d'un long processus itératif. Cette difficulté est fort bien exprimée par Rand Hindi, le CEO de la start-up SNIPS, dans son article *Will Artificial Intelligence Be Illegal in Europe Next Year ?*⁷⁰. S'il est parfaitement légitime pour l'Europe de protéger ses citoyens tout en rétablissant une forme d'équité qui s'applique à toutes les entreprises qui vendent leurs services en Europe, il faut réfléchir plus sur ce

69 Pour renforcer la compréhension du couplage données-modèles, on peut lire l'article de Ben Lorica, The current state of applied data science : <https://www.entrepreneur.com/article/298394>

70 Disponible sur le web : <https://www.entrepreneur.com/article/298394>

sujet complexe, qui dépasse le cadre de ce rapport. Il faut déplacer le concept de « finalité de la collecte » vers la « finalité de l'usage » et celui du « contrôle opérationnel » du processus de collecte (ce qui suppose un déterminisme qui est antagoniste avec certaines approches présentées dans ce rapport) vers **l'auditabilité** du processus de traitement des données (cf. chapitre V).

IL RESTE DES VRAIES BARRIÈRES SCIENTIFIQUES

Il reste de nombreuses questions difficiles à résoudre, sur le chemin d'une IA plus générale. Ces sujets sont traités plus en détail dans le rapport France IA [10] ou dans le rapport INRIA [13]⁷¹. Nous allons ici nous contenter de souligner les quatre points qui sont revenus le plus fréquemment durant les auditions de nos experts :

- avec les nouvelles méthodes issues de l'apprentissage, telles que le *Deep Learning*, il est souvent difficile de demander des explications au « système intelligent ». Ce n'est pas toujours nécessaire, lorsque la réponse ou l'action se situent dans des cadres très délimités. Par exemple, la société suisse Adhoco utilise l'apprentissage pour la domotique, dans un cadre d'interaction qui ne nécessite aucune explication. En revanche, certains domaines exigent des explications, comme l'utilisation de l'IA en droit, et d'autres sont plus efficaces lorsque les explications sont disponibles, comme la recommandation. Par exemple, le moteur de recommandations de contenus Jinni fait de sa capacité à expliquer ses recommandations un avantage compétitif. Il y a plusieurs pistes pour produire ces explications : soit par une approche hybride qui combine le symbolique et le neuronal, soit

71 Il faut d'ailleurs souligner la distance entre ce qui se fait dans un laboratoire de recherche et son application robuste plus tard dans une utilisation régulière : une partie de la confusion indiquée un peu plus haut vient du fait que la presse présente souvent des expériences de laboratoires comme des réalisations.

en entraînant le réseau de neurones à produire des arguments en même temps que sa décision⁷² ;

- de façon quelque peu similaire, les mêmes systèmes issus du *Deep Learning* sont pour l’instant incapables d’évaluer ou de qualifier leurs erreurs. Cela se traduit par des « erreurs spectaculaires », là où un humain moins précis ferait plus d’erreurs, mais moins importantes ou moins visibles. Certains experts pensent qu’il s’agit d’un problème fondamental pour l’application de l’IA dans des domaines où la sécurité est un prérequis intransigeant. La nature statistique des méthodes d’apprentissage — y compris les réseaux de neurones — se prête bien à la combinaison d’une réponse et d’un degré de confiance, mais cela demande des jeux de données plus étendus — dans l’esprit de l’*adversarial training* que nous avons évoqué plus haut — et des temps d’apprentissages encore plus long ;
- un des problèmes les plus difficiles et les plus anciens de l’IA est la capacité à intégrer « le raisonnement de sens commun »⁷³. Le sens commun est ce qui nous permet d’exercer une intelligence « large » et non pas « étroite » en convoquant un « contexte » autour de la question à traiter. C’est fondamental pour la compréhension du langage naturel pour laquelle la majorité des connaissances nécessaires sont implicites. C’est ce qui explique pourquoi Ray Kurzweil — comme indiqué précédemment — est très prudent sur la date à laquelle les *chatbots* pourront tenir des conversations « de généralistes ». Le sens commun nous permet d’éliminer toutes sortes de « faux positifs » dans les problèmes de reconnaissance de son ou d’image ; son absence est une des causes des « erreurs spectaculaires » des algorithmes de *Deep Learning* (que l’on peut constater avec les algorithmes de traduction, qui sont devenus très performants, mais continuent à nous faire sourire) ;

72 Ce thème général de recherche est qualifié d’*explainable artificial intelligence* — un des programmes DARPA. Voir par exemple : <https://www.darpa.mil/program/explainable-artificial-intelligence>

73 À ce sujet, Yann LeCun a déclaré dans le Wall Street Journal en 2016 : “*Teaching common sense to software is more than just a technical question—it’s a fundamental scientific and mathematical challenge that could take decades to solve. And until then, our machines can never be truly intelligent*”.

- la frontière la plus active entre le monde de la recherche et celui des applications est le passage de l'apprentissage supervisé (celui qui est utilisé aujourd'hui pour la majorité des applications) et l'apprentissage non-supervisé (qui existe dans nos laboratoires, mais n'est pas encore prêt pour la généralisation). La barrière n'est pas dans la non-supervision — il existe des approches non-supervisées en analyse de données depuis longtemps — mais dans la capacité à ouvrir le champ d'exploration (passer de « très étroit » à « moins étroit »). On retrouve ici le domaine des approches « génératives » évoquées dans le chapitre II.

IL EST DIFFICILE DE PRÉDIRE CE QUE L'ON NE COMPREND PAS

La confusion autour de l'Intelligence artificielle est renforcée par un discours volontariste de prospective qui affirme que les verrous seront vite débloqués et que les progrès continueront à un rythme exponentiel. C'est ce même discours volontariste qui justifie de porter le débat sur les conséquences d'une future « intelligence artificielle générale », au lieu de s'intéresser aux conséquences de l'intelligence artificielle étroite qui est en train de se répandre rapidement. Sur le thème prospectif des conséquences de l'IA forte, nous renvoyons le lecteur à l'excellent livre de Max Tegmark, *Life 3.0 : Being Human in the Age of Artificial Intelligence*⁷⁴. La prospective suppose que les différents verrous que nous venons d'évoquer seront levés dans les années qui viennent, en particulier grâce à la poursuite de la Loi de Moore. C'est la base de l'argument de la Singularité [25]. Nous allons revenir sur ces deux points, mais il faut retenir la thèse centrale de Max Tegmark, selon laquelle, en l'état actuel de nos connaissances, il n'est pas possible d'affirmer que ces verrous seront levés et que nous saurons produire une AGI, mais il n'est pas possible d'affirmer le contraire : « *We clearly don't*

74 Ce livre propose une excellente synthèse des questions qui se poseront alors, de l'état de nos connaissances par rapport à ces questions et de la diversité des avis de la communauté des experts.

know, and need to be open to the possibility that the answer may be “never”. However, a basic message of this chapter is that we also need to consider the possibility that it will happen, perhaps even in our lifetime ». Une partie des propos prospectifs alimentent la confusion entre la réalité et la science-fiction, ce qui produit par effet rebond des rapports ou des livres dans lesquels d'autres experts expliquent que « cela n'arrivera jamais ». Comme le souligne Max Tegmark, la réalité est que nous n'en savons rien⁷⁵.

Il est très difficile de faire des prévisions sur la continuation de la Loi de Moore, mais il existe beaucoup de signaux qui font penser à une inflexion. La loi de Moore spécifie que la densité des transistors dans les circuits intégrés au silicium à hautes performances double tous les 2 ans. Cette évolution est liée à la réduction de la largeur de grille des transistors élémentaires, qui atteint 10 nm en 2017 et pourrait se poursuivre avec des grilles de 7 nm et même 5 nm dans la décennie à venir, mais il semble difficile d'aller beaucoup plus loin sans se heurter à des courants de fuites et des difficultés propres à l'échelle atomique. Simultanément, la puissance de calcul des circuits a progressé en profitant de la miniaturisation, d'une part avec l'augmentation du nombre de composants et, d'autre part avec la montée en fréquence associée. Pour l'avenir, le rythme des gains en puissance de calcul pourrait ralentir, pour des raisons techniques et économiques⁷⁶. D'abord, la montée en fréquence s'est trouvée freinée, notamment par la capacité de dissipation du silicium, c'est le mur de la chaleur qui correspond à quelques watt/cm² en l'absence de radiateur et ventilateur, malgré les réductions des tensions de fonctionnement appliquées aux circuits. En fait, les fréquences d'horloge plafonnent autour de 4 GHz depuis une dizaine d'années. L'expérience semble montrer qu'il n'est pas raisonnable d'un point

////////////////////////////////////

75 Comme le remarque un des membres de la commission les plus experts sur ce sujet, cela signifie qu'il faut également rester modeste sur les recommandations proposées dans ce rapport : « Personne n'a réussi jusque-là à prédire les progrès de l'IA ou le pourquoi de ces progrès ».

76 Le rapport industriel ITRS-2015 (*International Technology Roadmap for Semiconductors*), publié en juillet 2016, prévoit que la loi de Moore restera valable jusqu'en 2021, mais que la miniaturisation au-delà de 10 nm ne sera plus souhaitable économiquement.

de vue pratique de dépasser les 50 W par circuit. Ensuite, les difficultés de réalisation sont devenues telles que la question de la justification économique de la poursuite de la miniaturisation est posée. On observe le passage vers des architectures silicium plus spécialisées, depuis les GPU jusqu'au ASIC. Les GPU sont plus efficaces en calcul par watt que les processeurs CPU traditionnels. C'est encore plus vrai pour les circuits spécialisés — on pense aux ASIC spécialisés pour le calcul de la blockchain ou aux puces neuromorphiques. On a vu apparaître les TPU (*TensorFlow Processing Unit*), mis en avant par Google, qui sont des ASIC spécialisés pour l'exécution efficace des réseaux neuronaux profonds (en particulier *TensorFlow*). Ces chipsets spécialisés pour implémenter des algorithmes de réseaux neuronaux, comme la puce *TrueNorth* qui encode 1 million de neurones et 256 millions de synapses, sont beaucoup plus économes en énergie pour une quantité de calcul et d'apprentissage donnée⁷⁷. Finalement, il est difficile de dire ce que deviendra la loi de Moore⁷⁸ (certains lui prédisent une fin comparable à celle du fleuve Okavango qui se perd dans les sables du désert du Kalahari). Simplement, on peut s'attendre à un ralentissement du rythme d'augmentation de la puissance de calcul des circuits au-delà de 2020.

Les capacités de l'IA sont bien plus difficiles à prévoir que l'augmentation de la puissance de calcul, mais cet argument joue dans les deux sens. Sans rentrer dans le détail de l'argument de la Singularité (voir [25] [19] [16]), son principe fondamental est que la croissance continue et exponentielle de la puissance de calcul nous conduit forcément à un point critique où l'intelligence artificielle dépasse l'intelligence humaine. Nous venons de voir que cette croissance exponentielle n'est pas acquise, mais il est encore plus clair qu'il n'existe pas de dépendance simple entre les performances de raisonnement d'une intelligence

////////////////////////////////////

77 Voir également les annonces des puces spécialisées pour les smartphones, comme celle de Huawei : <http://www.zdnet.com/article/huawei-unveils-kirin-970-chipset-with-ai/>

78 La prospective sur l'utilisation des ordinateurs quantiques en tant que génération suivante pour supporter les progrès de l'intelligence artificielle dépasse le cadre de ce rapport. Lire par exemple l'article suivant sur la collaboration entre le MIT et IBM : <https://www.technologyreview.com/s/608810/ibm-and-mit-bet-that-materials-and-quantum-advances-will-supercharge-ai/>

artificielle (sa capacité à passer d'un domaine étroit à un domaine général) et la puissance de calcul disponible. Il y a ici un consensus fort des différents experts que nous avons écoutés : « la puissance de calcul ne suffit pas ». En revanche, de façon symétrique, la difficulté à prévoir signifie qu'il faut se méfier des discours réducteurs sur les « limites théoriques » de l'IA. Il s'est produit une révolution dans le domaine de la perception depuis 2010 grâce aux réseaux neuronaux profonds, et il est raisonnable de penser que nous n'en sommes qu'au début des impacts à long terme de ces progrès⁷⁹. Ces progrès en perception vont donner naissance par hybridation à des systèmes qui combinent perception, traitement du langage naturel et raisonnement symbolique. L'accès à de très grandes bases d'expériences de perception va former une nouvelle fondation pour le raisonnement de sens commun. Il serait vain de vouloir prédire ce qui va émerger, mais il est clair que les pistes pour lever les verrous précédemment évoqués sont multiples et ne supposent pas nécessairement la poursuite de la Loi de Moore sur le même rythme.

79 Pour faire l'expérience en direct de ces progrès, il faut utiliser l'application produite par Microsoft sur un iPhone à destination des non-voyants pour identifier les objets quotidiens dans une scène.

Favoriser le développement de l'IA dans les entreprises

COMPRENDRE LES CONDITIONS DE SUCCÈS

Pour profiter pleinement des avancées de l'Intelligence artificielle pour mieux traiter ses problèmes métier, l'entreprise doit travailler un ensemble de conditions favorables, pour faire « émerger » les succès. Le terme d'« émergence » traduit une approche *IA-ready* : se mettre dans les conditions pour profiter à la fois des opportunités technologiques qui arrivent en flux continu de l'extérieur (des nouveaux algorithmes, des API de nouveaux services, des nouveaux capteurs, etc.) et des opportunités métiers que l'ensemble des acteurs de l'entreprise peut identifier dès lors que les données nécessaires ont été collectées. Voici un petit résumé des conditions que l'entreprise doit réunir pour être *IA-ready* (prête pour l'Intelligence artificielle) :

- la première conséquence du chapitre III est qu'il existe des opportunités partout, mais qu'il est difficile de savoir à l'avance ce qui va apporter de la valeur ;
- Il ressort de ce que nous avons écrit plus haut que la première étape de la stratégie est la collecte et la curation de données ;

- les outils sont multiples, facilement accessibles et se renouvellent constamment. L'objectif des entreprises est donc de construire un environnement logiciel qui favorise l'intégration continue de ces nouveaux outils ;
- comme expliqué précédemment, l'apprentissage et l'utilisation de l'IA requièrent une démarche scientifique et un peu de rigueur statistique pour éviter de tomber dans les pièges des « résultats non significatifs » ;
- pour finir, il s'agit en premier de lieux de pratiques, dont une partie importante de la valeur est liée à l'expérience, qui demande du temps.

La pyramide de la figure 8 est une représentation simplifiée, à la manière de la « Pyramide de Maslow » des conditions que nous venons d'évoquer.

À cause de la complexité des causes favorables, et de la lenteur des acquisitions des compétences et des processus d'analyse des données, le pilotage doit suivre une approche de type *Affordable Loss* empruntée de l'innovation numérique, plutôt que de chercher à construire des *Business cases* rigoureux. L'incertitude de la création de valeur lorsqu'on applique des approches d'intelligence artificielle n'est pas sans rappeler l'incertitude dans la création de nouveaux services numérique, due à la non-linéarité des processus d'adoption par les utilisateurs [6]. En effet, la création de valeur passe le plus souvent par la mise en place d'un cercle vertueux « IA → insights → plus d'usage → plus de données → meilleure IA », un processus non-linéaire qui a les mêmes propriétés chaotiques, et est donc difficile à prévoir (cf. l'argument du co-développement du modèle métier et de la plate-forme intelligente de support). En conséquence, il est judicieux d'emprunter l'approche *Affordable loss* du monde numérique. Plutôt que de construire un *business case* rigoureux qui suppose de pouvoir évaluer à l'avance la création de valeur, il est préférable de distinguer un investissement stratégique de « capacité » (architecture de données, collecte des jeux) et des projets d'investigation légers, dont le financement est déterminé par la somme maximale que l'entreprise peut accepter de « jouer pour voir ». Pour fonction-

ner, cette approche suppose une capacité agile à fonctionner avec un budget fixe et une capacité à abandonner sans regret tout ce qui ne fonctionne pas⁸⁰.

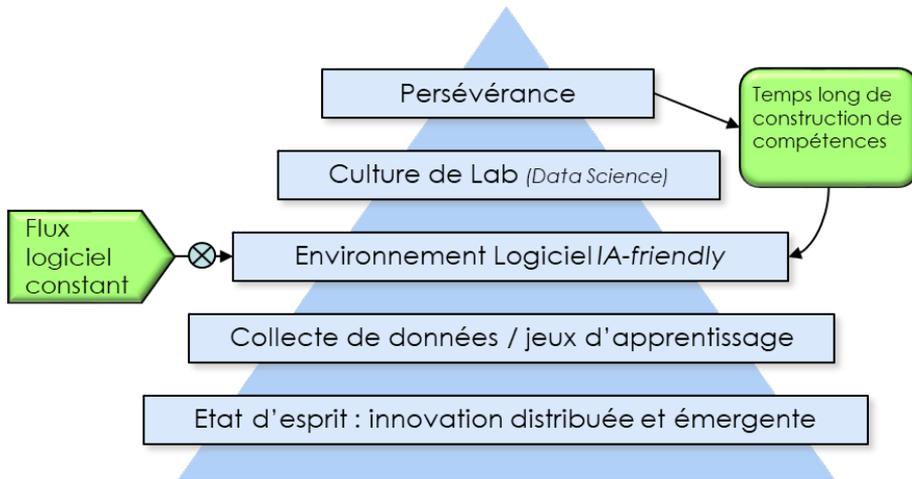


Figure 8 : Pyramide hiérarchique des conditions de mise en place d'une démarche IA

TRAQUER LES IDÉES REÇUES

Comme il y a énormément d'incertitude sur ce qu'il sera possible de faire, la non-connaissance est remplacée par des idées toutes faites. Il est important, pour construire une stratégie sur l'utilisation de l'Intelligence artificielle d'éviter deux écueils : l'exagération de ce qui est possible qui conduit à des attentes déçues comme nous l'avons vu au chapitre III, ainsi que les « fausses certitudes » sur ce que l'IA ne pourra jamais faire, qui conduit à des stratégies attentistes qui risquent de se heurter à cet « avantage au premier joueur ». Curieusement⁸¹,

80 Plus qu'un *fail fast*, il s'agit d' « apprendre en itérant aussi vite que possible et en s'arrêtant si l'on dépasse un montant prédéterminé de dépenses ».

81 « Curieusement » ou « classiquement », puisqu'on doit à Kant cette citation redoutable : « *On mesure l'intelligence d'une personne à la quantité d'incertitude qu'il est capable de supporter* ». À méditer dans le cadre de l'intelligence artificielle, puisque les programmes sont aussi capables

l'incertitude que nous avons décrite dans le chapitre précédent est « inconfortable » pour de nombreux acteurs qui préfèrent une représentation fautive, mais simplifiée, ce qui peut nous empêcher de considérer l'ensemble des branches possibles du futur de l'IA. À titre d'illustration, nous avons étudié un rapport du *World economic forum* [28] — au demeurant fort intéressant — qui traite du futur de l'Intelligence artificielle. Notre analyse souligne quelques raccourcis dont il faut se prévenir :

- penser que l'IA est passée des moteurs de règles hier à l'apprentissage aujourd'hui ;
- mettre sur le même plan la recommandation automatique à partir d'inputs de la chaîne de valeur (très court terme) et l'utilisation l'apprentissage non-supervisé pour construire des chatbots généralistes (long terme) ;
- considérer que l'étape « les moteurs de recherche fourniront des (vraies) réponses au lieu d'une liste de résultat » est très loin dans l'horizon (alors que la synthèse de réponse en langage naturelle commence à bien fonctionner pour de nombreux domaines — cf. les robots qui écrivent des articles) ;
- l'affirmation « l'intelligence artificielle générale requiert des capacités différentes de calcul non-déterministe » est totalement non prouvée.

Il ne s'agit pas de chercher un consensus sur ce qui sera possible — alors que les experts ont des avis fortement divergents —, mais de traiter ce sujet avec une approche de gestion des risques. Une position réaliste consiste à ne pas croire les promesses de ceux qui extrapolent les progrès passés, mais à ne pas croire non plus ceux qui « savent ce qui est impossible ». Une approche fondée sur le risque consiste à combiner une politique agressive de déploiement de l'IA sous forme de petits pas autour de ce qui est possible aujourd'hui ou dans le futur proche, avec une approche stratégique défensive consistant à observer et à considérer l'avancée rapide des capacités de l'IA comme un risque (par exemple de disruption par un nouveau compétiteur). En termes d'analyse de risque (probabilité x impact), le futur proche fournit des impacts mesurés

de raisonner sur l'incertain.

à haute probabilité et le futur plus lointain des probabilités faibles (en l'état actuel de notre savoir), mais des impacts importants (en automatisation ou performance). Les deux doivent faire partie du portefeuille stratégique de l'IA, mais en restant lucide sur les différentes catégories. Comme nous l'avons dit à la fin du chapitre III, le progrès n'est pas continu, mais sous forme de saccades causées par un enchaînement complexe de progrès intermédiaires. Les progrès spectaculaires des dernières années en termes de perception vont avoir des conséquences multiples, qui vont bien plus loin que les véhicules autonomes ou la surveillance automatisée.

CONSTRUIRE UNE CULTURE PROPRE À CHAQUE DOMAINE D'ACTIVITÉ

La plupart des entreprises qui utilisent l'IA avec succès développent progressivement une approche qui est propre à leur métier, leur problématique et leurs données. Un message important de nos experts est que l'IA évolue en ce moment vers la spécialisation, c'est-à-dire le développement de multiples approches dépendantes du domaine d'application. C'est la conséquence directe de l'état de l'art aujourd'hui qui est une capacité étroite à résoudre des problèmes précis. Cela conduit naturellement à la prolifération et à la diversification. Pour les entreprises cela signifie qu'il n'y a pas « d'approche type » : les choix de modèles, d'algorithmes et de protocoles d'apprentissage et d'intégration sont propres aux domaines et aux types de données collectées. La coévolution entre modélisation métier et développement de solutions « intelligentes » rend la spécialisation par entreprise encore plus importante. « Il n'y a pas d'intelligence artificielle universelle, les progrès se trouvent aux interfaces entre les différentes approches et les domaines métiers » nous a expliqué Francis Bach. Une des conséquences de cette diversification est qu'elle représente des besoins importants en investissement. Pour la majorité de nos experts, la France n'investit pas assez dans le déploiement des méthodes dans l'ensemble des domaines d'activité.

Il faut considérer l'ensemble des acteurs et des outils comme un écosystème auquel il convient de contribuer. Parmi de nombreux exemples similaires, le programme de la *Singularity University* cite l'exemple de Pinterest qui a construit une collaboration rapide avec une petite équipe de Berkeley pour améliorer ses méthodes de classification thématique d'images : « *Pinterest faced this : deep learning to allow people to search on the interesting part of their pictures. Three people from their company, one professor, two students. It was wildly successful and took six months.* » Apprendre à travailler avec l'écosystème demande un équilibre : il ne faut pas « tout faire en interne » mais pas non plus tout externaliser, comme le souligne Francis Bach. L'entreprise doit avoir des gens bien formés, qui peuvent parler aux autres acteurs : les académiques, les spécialistes des start-up ou les développeurs de plates-formes. Il faut savoir nouer des relations informelles plus que contractuelles. Cette approche d'écosystème est la raison pour laquelle les grands acteurs publient leur solution en *open source*. C'est aussi pour cela qu'IBM a annoncé en 2014 qu'au moment où il investissait 1 milliard sur *Watson*, il constituait un fond de 100 millions d'investissement *garage level* dans des start-up qui s'appuient sur *Watson* (conséquence de la diversification que nous avons évoquée plus haut).

Les recommandations de ce rapport ne sont pas très différentes de celles faites par Neil Jacobstein lors de son intervention à la *Singularity University*, qui indique qu'il faut :

- investir sans attendre (en fonction de ses moyens propres à chaque entreprise) ;
- utiliser la boîte à outils d'algorithmes disponibles⁸² ;
- apprendre à utiliser les hardwares spécialisés (GPU, TPU, ASIC) ;

////////////////////////////////////
 82 Après la phase d'obsession sur le Deep Learning, on peut constater une position plus équilibrée dans l'article récent de la Harvard Business Review (<https://hbr.org/2017/08/a-survey-of-3000-executives-reveals-how-businesses-succeed-with-ai>) qui insiste sur une approche de « portefeuille » qui exploite l'ensemble de la boîte à outils — utiliser les méthodes éprouvées à court terme, expérimenter les approches nouvelles edge stuff à moyen terme et construire des partenariats de recherche pour le long terme, avec comme première priorité l'infrastructure données / digitale (*digital capabilities before AI*).

- faire l'inventaire et rendre disponible l'ensemble des sources de données de l'entreprise ;
- utiliser des approches collaboratives pour travailler avec des *data scientists* externes.

Il est intéressant de noter qu'un des membres de la commission TIC a pu entendre les mêmes recommandations pratiques que celles présentées dans ce rapport, de la part de jeunes programmeurs en Inde à Bangalore — recommandations simplement issues d'une pratique des outils et méthodes présentés au début du chapitre II. Il est rassurant de constater que les recommandations des experts sont confirmées par l'observation pratique, mais il faut surtout noter que la révolution de l'application massive de ces technologies est en marche ! Rappelons également ici l'importance de la capacité de calcul qui détermine la vitesse du cycle itératif de méta-apprentissage. Pouvoir accéder à des fermes de machines spécialisées (comme des GPU ou TPU) est un formidable accélérateur d'innovation et de compétitivité.

Puisque la stratégie de l'émergence de l'IA commence par la constitution de corpus de données métiers, l'infrastructure de collecte et partage de données joue un rôle essentiel. Du point de vue des systèmes d'information, il y a trois domaines : celui de la collecte, interne et externe, celui du stockage, traitement et manipulation, à froid comme à chaud⁸³, et celui de la création de valeur, lorsqu'on applique les méthodes présentées dans ce rapport. Les recommandations de tous les experts, y compris Neil Jacobstein ou ceux de la *Harvard Business Review*, sont de construire le deuxième tiers sans attendre, de façon globale en évitant la formation de tout silo et en s'attachant à l'ouverture — la capacité à intégrer des sources de données externes. Il s'agit d'une stratégie hybride, qui combine une infrastructure de collecte et mise à disposition de données robuste

83 Les informaticiens distinguent les traitements à chaud "en temps réel" sur les flux de données — en particulier les événements — des traitements analytiques à froid obtenus par "fouille des données". Du point de vue fonctionnel il y a une complète synergie — l'intelligence artificielle s'applique aux deux — mais d'un point de vue technologique, il est préférable de distinguer des plates-formes différentes.

et industrielle — un investissement stratégique pour l'entreprise — et plusieurs satellites de valorisations de ces données pour lesquels la diversité, la prise de risque et l'agilité sont des atouts essentiels (cf. chapitre IV). Les acteurs les plus avancés font des choix conservateurs pour construire ce cœur stable et préfèrent choisir des solutions éprouvées comme le rapporte Cecil Dijoux dans son livre *Que signifie l'avènement du numérique* [3]. L'audace technologique se concentre « sur les bords », sur les plates-formes *Data labs* avec lesquelles les techniques d'Intelligence artificielle et d'apprentissage sont mises en place. L'objectif de l'infrastructure analytique de donnée est précisément de construire l'agilité qui permet d'utiliser les nouvelles plates-formes — en particulier sur le cloud, proposées « *as a service* » — des acteurs les plus innovants / puissants de l'écosystème (on pense bien sûr, par exemple, à Google).

Recommandations

SOUTENIR L'EFFORT DE PROMOTION ENGAGÉ PAR « FRANCE IA »

La France et l'Europe sont en retard par rapport aux États-Unis et à la Chine.

Les recommandations de nos experts sont en ligne avec ce qui a été rappelé au chapitre précédent. Il existe un besoin important en investissements, qu'il s'agisse de ressources de calcul, de données ou de recherche, et dans ce domaine les États-Unis et la Chine sont en avance⁸⁴. Il est probablement nécessaire de construire un « plan IA » à la fois à l'échelle européenne et nationale, une sorte de « programme Apollo » pour reprendre les termes d'un expert. Il faut un plan d'investissement « multi-échelle » à cause de la diversité des applications et des domaines que nous avons évoqués dans la section précédente. Il existe un double enjeu par rapport à l'intelligence artificielle : un enjeu économique, celui de la compétitivité de nos entreprises — cf. le rapport de la Maison-Blanche — et un enjeu de souveraineté⁸⁵. Comme le rapportent un certain nombre de cher-

84 En ce qui concerne la Chine, il faut modérer ce propos puisque nous avons vu qu'il faut accumuler de l'expérience et pas seulement des investissements ou des algorithmes. Ce point de vue intéressant est développé dans l'article *Is China really that far ahead in AI ? Research says no* : <http://technode.com/2017/08/24/is-china-really-that-far-ahead-in-ai-research-says-no/>

85 Rappelons ici la citation de Vladimir Poutine : « *L'intelligence artificielle est l'avenir, pas seulement de la Russie, mais de toute l'humanité. Elle présente des opportunités colossales, mais également des menaces qui sont difficiles à prévoir. Quiconque devient le leader de ce secteur deviendra le maître du monde* »

cheurs français — au rang desquels on trouve Stéphane Grumbach — il ne faut pas « laisser Google ubériser l'État ». Aujourd'hui, il est clair que les avancées et la recherche sont fortement tirées par des financements privés.

Le financement de la recherche, tant par son mode que son volume, n'est pas suffisant. Nous avons recueilli de nombreuses inquiétudes des experts, à la fois sur le financement tant national qu'europpéen. Le modèle européen se concentre sur des grands projets, pour favoriser une concentration des moyens. Cette recherche « chère » n'est pas toujours efficace. Cette approche conduit à faire des impasses dans le temps et dans l'espace des thèmes. Dans le temps, cette vision orientée sur les projets et les résultats visibles conduit à un manque de continuité dans les thèmes. Du point de vue de la couverture des domaines, comme il est devenu « détestable de saupoudrer », nous avons, de fait, perdu le principe de diversification de la recherche. Face à la diversité des défis exposée au chapitre précédent, l'Europe et la France auront des difficultés à soutenir la recherche, à la fois fondamentale et appliquée. Il y a par ailleurs, cela a été dit, un formidable défi d'intégration d'approches différentes, ce qui demande à la fois un renforcement des investissements, une intensification de la dimension technologique et une véritable diversification des domaines supportés par le financement de la recherche.

Il faut constituer des lieux d'excellence qui attirent et concentrent des « masses critiques » de chercheurs et de praticiens. Nombre de nos experts ont vu « leurs meilleurs étudiants » attirés par les grands acteurs américains, à cause de l'opportunité de rejoindre des équipes d'excellence et de bénéficier de salaires souvent au moins dix fois plus élevés. Il y a bien sûr une question de moyens, que nous venons d'évoquer — en particulier l'accès à des plates-formes qui combinent des très grands volumes de données et des très gros moyens de calculs modernes —, mais il y a aussi la valeur attractive de la concentration de talents dans le domaine. Nous sommes dans une phase de création et de développement intense, pendant laquelle la « fertilisation croisée » des méthodes

et des approches est particulièrement pertinente... et passionnante pour les jeunes chercheurs. Nous avons donc besoin de créer des lieux d'excellence en Europe, privés et publics (et les combinant), qui puissent rivaliser avec Facebook et Google pour y « recruter les meilleurs » et créer cette masse critique⁸⁶.

La question du « retard » de la France est complexe et fait donc débat. Non seulement, comme le souligne le rapport France IA [10], nous avons une excellente communauté de recherche française, à la fois sur les domaines cœurs (par exemple, les réseaux neuronaux) et sur les domaines connexes (par exemple, la démonstration de théorèmes) et la contribution de la France dans le monde est très significative. Comme cela est souvent remarqué, les meilleurs contributeurs aux États-Unis sont souvent des Français expatriés et les logiciels *open source* font la part belle aux contributions françaises. La bibliothèque *Scikit-learn* a été une des plus influentes dans le monde de l'apprentissage. L'affirmation que la France est en retard consiste à prendre plus de recul et juger la taille et l'impact des acteurs qui utilisent l'IA (par exemple, Criteo), ainsi que les moyens de calcul, les volumes de données et la taille des équipes assemblées pour développer les applications de demain. Même si certaines grandes entreprises françaises sont très avancées dans l'utilisation des méthodes décrites dans ce rapport pour leur propre domaine d'activité, la France, en tant qu'écosystème industriel global, présente un certain retard.

La recommandation la plus importante de l'Académie des technologies est de favoriser l'émergence d'une certification des « processus intelligents et / ou autonomes » qui utilisent les méthodes de l'Intelligence artificielle. Comme cela a été dit plus haut, c'est bien l'ensemble du processus décrit au chapitre II qui doit être certifié, depuis la collecte de la donnée jusqu'à l'usage en passant

86 Ce point ne fait pas consensus parmi tous les experts, puisque la création de centres importants est souvent synonyme de concentration sur les approches et les technologies. Il faudrait donc combiner la massification et la diversité. De façon plus générale, il y a en revanche un consensus sur le besoin pour la France d'embaucher et de développer plus de talents autour des disciplines de l'Intelligence artificielle, dans le monde de la recherche comme celui de l'entreprise.

par les protocoles d'apprentissage. Cette certification suppose la construction de « centres d'essais et de validation » propres à chaque domaine qui rassemblent des jeux de données, des moyens de calculs et des compétences métiers propres au domaine⁸⁷. L'industrie de l'automobile en Allemagne donne l'exemple d'une telle démarche, mais des réflexions semblables sont en cours en France. Comme le fonctionnement des algorithmes d'apprentissage n'est pas forcément déterministe et que les données collectées ne sont pas forcément représentatives de tous les cas d'usage, cette certification s'appuie sur le concept de finalité **auditable** du processus [14]. Il faut encourager les différentes parties prenantes, des industriels à la communauté scientifique en incluant les utilisateurs, à développer des normes auditables qui serviront à construire les systèmes intelligents et autonomes de demain.

TRAITER LES QUESTIONS ÉTHIQUES DE L'IA LE PLUS TÔT POSSIBLE

L'intelligence artificielle pose des questions éthiques sur le contrôle et la responsabilité qui doivent être traitées. Le groupe de travail de l'Académie a reçu deux représentants de l'AFCIA (Association française contre l'Intelligence artificielle), Cédric Sauviat et Marie David, et le consensus s'est rapidement établi sur le fait que le développement et l'usage de l'Intelligence artificielle posent de graves questions. Ces questions sont en premier lieu des questions de contrôle et de maîtrise des risques et, dans un deuxième temps, des questions éthiques sur l'impact de l'usage de l'IA sur la société — quelles modifications dans nos vies et nos sociétés, et quels modes de gouvernance pour accepter ces changements ? Dans le prolongement du travail de l'Académie sur l'acceptabilité sociétale des TIC, il nous apparaît que les développements des techniques, de

87 Cette recommandation provient de la note de l'Académie des technologies à l'attention du Président de la République pour le développement d'une informatique prédictive et prescriptive (HPC-IA), envoyée le 13 juin 2017. On trouve une recommandation semblable dans le rapport du CGE IET [14] ainsi que dans les recommandations de l'IRT SystemX pour la mission Villani : construire des plates-formes technologiques collaboratives par domaine et travailler sur la validation de systèmes critiques intégrant l'IA.

leur usage et de leur compréhension doivent être harmonieusement mêlés. Pour reprendre une des interrogations de nos intervenants : « on ne peut pas à la fois prédire l'émergence d'une conscience artificielle et ne pas se poser des questions éthiques sur l'usage et le traitement des robots ».

Il y a plusieurs débats autour de l'intelligence artificielle qui correspondent à des horizons de temps différents. Nous allons ici brièvement évoquer les quatre débats les plus importants⁸⁸, dans l'ordre de plausibilité du risque — au fur et à mesure que les possibilités de l'IA augmentent :

- la question de l'équité des « décisions artificielles » — des recommandations ou décisions prises par des algorithmes — est entière et n'a rien de théorique. Les algorithmes qui sont construits par apprentissage à partir d'un jeu de donnée héritent du double biais de modélisation du concepteur et de collecte des données. Le rapport du CGIET cité au chapitre II donne des exemples et des explications très convaincantes : il existe bien une vraie difficulté éthique et sociétale (du point de vue de ce qui est juste et de ce qui est acceptable). Comme cela a été dit plus haut, l'Académie fait sienne les préconisations de ce rapport, sur le besoin de développer des méthodes de test, d'audit et de certification de l'équité des algorithmes qui utilisent l'intelligence artificielle et l'apprentissage. L'apprentissage à partir d'une grande masse de données produit facilement des « règles conservatrices » qui fonctionnent pour le plus grand nombre au détriment des minorités et des personnes en situation de fragilité sociale⁸⁹ ;

88 Pour une analyse plus détaillée des questions d'éthique liées aux systèmes qui s'appuient sur l'apprentissage, il faut lire le rapport *Éthique de la recherche en apprentissage machine* du CERN (Commission de réflexion sur l'éthique de la recherche créée en 2012 par l'alliance Allistene) : http://cerna-ethics-allistene.org/digitalAssets/52/52472_CERNA_Ethique_de_la_recherche_en_apprentissage_machine.pdf

89 Cette question des biais des systèmes automatiques est profondément complexe et pourtant d'une grande actualité. Il faut lire, comme le recommande Cédric Villani, le livre de Cathy O'Neil *Weapons of Math Destruction* [2] qui donne de nombreux exemples dans lesquels l'instauration d'un principe algorithmique de décision crée des biais importants et pas forcément faciles à détecter.

- la question de l'acceptabilité d'une « décision automatique », de la part d'une intelligence artificielle, est plus large. Il s'agit pour la société de décider quand il est acceptable de remettre des décisions à un big brother. L'AFCIA s'interroge à juste titre sur les processus de gouvernance qui vont permettre à la société d'exercer ses choix, par opposition au développement non contrôlé d'un technosystème « mû par ce qui est possible » ou le développement commercial tiré « par ce qui est rentable ». Les sociétés démocratiques disposent des outils pour traiter ces questions de gouvernance, du moment que la transparence est assurée et que les pouvoirs publics assurent l'instruction des débats. Cette affirmation ne tranche pas la question pour ceux qui, comme l'AFCIA, rejettent l'objectif même de l'intelligence artificielle et répondent de façon négative à cette deuxième question ;
- le troisième débat qui s'installe avec l'arrivée de toute technologie qui permet l'automatisation des tâches est celui de la perte ou le déplacement des emplois. Ce débat dépasse clairement le cadre de ce rapport. Les arguments qui ont été développés dans les sections précédentes montrent que la question de l'impact de l'IA sur l'emploi est éminemment pertinente, mais que la prospective est difficile. L'Académie a travaillé depuis un certain temps sur ce sujet — par exemple à l'ordre du jour de notre séminaire de 2016 — qui requiert une approche transdisciplinaire qui dépasse le simple cadre de la commission TIC (collaboration avec des économistes, des sociologues et des spécialistes de l'industrie) ;
- le quatrième débat porte sur le contrôle des « algorithmes intelligents » et notre capacité à savoir et maîtriser ce qu'ils font. Dans le cadre futur d'une IA « générale », cette question est associée à un risque fort, puisque les capacités de cette « super intelligence » seraient grandes. cette question se pose aussi tout de suite, avec des algorithmes plus « étroits », mais qui sont déjà en train d'agir dans nos vies, depuis des choses simples comme la recommandation dans l'e-commerce (sujet qui est le cœur du rapport du CGIET) jusqu'à la conduite autonome des véhicules. Dans le cas où l'automatisation est socialement acceptable (question 2), la question est ensuite

de savoir si son fonctionnement est conforme aux intentions annoncées. Ici aussi, on retrouve des principes et exigences de transparence sur les intentions, associés à la mise en place de certification, ce qui suppose une certaine forme d'auditabilité. Il est clair que ce sujet est à la frontière de nos connaissances et compétences, et qu'il convient de développer les outils et méthodes de certification au fur et à mesure du développement de l'Intelligence artificielle. Il faut donc soutenir et encourager l'approche de *Responsible AI* qui vise à codévelopper simultanément les performances et l'acceptabilité éthique des algorithmes intelligents.

Nous avons décidé dans ce rapport de ne pas aborder les questions qui tournent autour des risques futurs d'une « super intelligence ». D'une part parce qu'il est très difficile d'évaluer la plausibilité de ce risque, et d'autre part parce que le développement mériterait un autre rapport tout aussi long, le sujet étant très complexe. Nous renvoyons le lecteur au livre de Max Tegmark déjà cité [19], qui propose à la fois un tour d'horizon très complet sur la diversité des points de vue et une analyse très structurée des questions qui restent ouvertes.

L'Académie des technologies porte un avis favorable aux recherches en Intelligence artificielle et son application dans la vie quotidienne des utilisateurs et des entreprises. Un des débats portés par l'AFCIA — parmi de nombreux autres — est celui du principe de précaution appliquée à l'intelligence artificielle, dès lors que le risque semblerait trop important. Il nous semble que le seul risque « trop important » est celui de la « super intelligence » qui reste complètement théorique. Les débats importants sont ceux du présent et du futur proche, pour lesquels nous avons les outils pour aborder les questions précédemment évoquées. L'intelligence artificielle n'est pas une discipline isolée qu'on pourrait séparer du reste de la science et des technologies de l'informatique. L'Académie des technologies considère que sa devise *Pour un progrès raisonné, choisi et partagé* s'applique naturellement aux travaux sur l'Intelligence artificielle, et que le progrès technologique est indispensable à la société pour faire face à

tous les défis que nous connaissons. Le statu quo n'est pas durable (au sens du développement durable) et il n'est pas clair qu'il soit souhaitable non plus. Dans le cas de l'Intelligence artificielle, les questions du contrôle et de la gouvernance de ce contrôle sont essentielles — et loin d'être résolues aujourd'hui — , mais les solutions à ces questions doivent être codéveloppées au fur et à mesure que la technologie et nos connaissances progressent.

PROMOUVOIR L'INTELLIGENCE ARTIFICIELLE DANS UNE PERSPECTIVE LARGE

Il ne faut pas « enfermer l'IA » dans une vision étroite et la séparer des autres disciplines informatiques auxquelles elle est liée. Ce rapport a exprimé deux idées essentielles. Premièrement l'IA n'est pas un but en soi, mais un outil qui peut servir à des fins multiples et qui est donc destiné à pénétrer la majorité des pratiques et des environnements, dans l'entreprise comme dans la société civile. Deuxièmement, l'IA n'est pas une technique spécifique et facilement identifiable, mais un ensemble large de méthodes qui ont vocation à s'intégrer dans l'ensemble des méthodes et outils informatiques. Nous reprenons ici une idée largement développée dans le rapport précédent sur le Big Data [1]. Dans un monde où les machines vont nous aider de mieux en mieux à trouver des réponses, les expertises de demain vont se déplacer vers d'autres questions — il restera toujours aussi important de savoir poser les bonnes questions — et vers d'autres défis de réalisation qui seront rendus possibles par cette collaboration homme-machine. L'opposition homme-machine, tirée par la racine de « l'imitation humaine » est l'arbre qui cache la forêt de l'« intelligence augmentée », celle de la collaboration entre l'homme et ces nouveaux outils cognitifs⁹⁰.

90 Il est plus que probable que ce qui est intéressant dans l'utilisation de l'intelligence artificielle n'est pas le fait de penser mieux que l'homme (la même chose, plus vite et avec moins d'erreurs), mais autrement. C'est le sujet de l'article célèbre de Kevin Kelly, *The Myth of Superhuman AI* paru dans Wired en Avril 2017, <https://www.wired.com/2017/04/the-myth-of-a-superhuman-ai/>

Il faut parler de l'intelligence artificielle et de ses succès de façon pragmatique pour participer aux conditions favorables à son développement.

Comme cela est souligné dans le rapport de l'OPECST [22], le discours public en France fait une large part à la question des conséquences de l'IA d'après-demain, ce qui est à la fois spéculatif et subjectif, tandis que le sujet des conséquences directes de l'IA aujourd'hui et demain est relégué en deuxième plan. Un de nos experts remarque que le sujet de l'IA dans les entreprises est vécu différemment en France qu'il ne l'est aux États-Unis à cause de notre plus grande « sensibilité sociale ». Les pouvoirs publics ont donc un rôle à jouer, celui de mettre en avant les progrès qui sont accomplis dès aujourd'hui grâce aux algorithmes qui utilisent les méthodes présentées dans ce rapport. Il est important de créer un contexte positif autour du progrès des technologies, qui unifie les perspectives des utilisateurs, de la société, des chercheurs, des entreprises et de la presse. Le rôle fondamental de la donnée — et de sa collecte — , qui a été souligné à de nombreuses reprises, signifie qu'il existe un écosystème complexe entre les chercheurs, les concepteurs, les industriels et les utilisateurs⁹¹. Pour fonctionner harmonieusement, compte tenu de cette complexité et des échelles de temps entre l'usage d'aujourd'hui et l'enrichissement des services de demain, il faut un alignement global collectif autour d'une vision partagée et pacifiée des ambitions de cette collaboration entre humains et machine. C'est très précisément ce qu'essaye de faire le rapport de la Maison-Blanche [29]⁹². Pour favoriser le développement compétitif de l'Intelligence artificielle en Europe, il ne suffit pas de donner plus de moyens aux chercheurs, il faut créer un écosystème de marché favorable autour du public utilisateur. Tim O'Reilly souligne à juste titre l'importance des utilisateurs dans la mise au point des systèmes intelligents

////////////////////

91 Par exemple, il nous semble intéressant de favoriser – financièrement et par la reconnaissance – la double rotation de chercheurs pour quelques années en entreprise, et des ingénieurs dans les organismes de recherche. INRIA a déjà démontré la faisabilité et la pertinence de cette approche, mais elle pourrait être grandement étendue.

92 Ce point est souligné par Laurent Alexandre dans son article du 17 Novembre 2017 dans les Echos. La réussite de l'écosystème de l'Intelligence artificielle, depuis les données jusqu'aux plates-formes et aux applications passe par des usages et une collecte de données qui reposent sur la confiance. Les efforts des think tanks américain visent à rassurer l'opinion publique.

modernes, ce qui implique que la confiance des utilisateurs et l'acceptabilité des technologies sont des ingrédients essentiels du succès de l'écosystème de l'Intelligence artificielle.

Dans le prolongement des recommandations de France IA, il est souhaitable de construire un cadre de référence pour le développement de l'IA en Europe.

L'Intelligence artificielle est un sujet « à la mode », il suffit de parcourir la presse, les réseaux sociaux ou le Web pour s'en rendre compte. Il est clair qu'il y a trop de variations dans la définition du vocabulaire associé, la compréhension du périmètre — c'est pour cela que ce rapport a insisté autant sur une vision large — ou sur les objectifs. Plusieurs de nos experts, dont Patrick Albert qui s'est également beaucoup investi dans le mouvement « France IA », militent pour la constitution d'un cadre de référence. Il conviendrait de reformuler les objectifs du domaine, de façon plus large que la « simple » imitation de l'intelligence humaine et en se détachant des racines historiques. La communication autour de la recherche en intelligence artificielle gagnerait à être articulée autour d'un cadre commun proposé et reconnu par la communauté des chercheurs. Un cadre de référence sur l'Intelligence artificielle permettrait d'y associer dès ses fondations la notion d'IA « responsable » en incluant les dimensions éthiques que nous avons trop brièvement abordées, les questions complexes du contrôle, de la certification et du test⁹³. Ce cadre de référence pourrait être associé à un observatoire de l'Intelligence artificielle, placé sous le contrôle d'une agence européenne, dont l'objectif serait de suivre et référencer les pratiques et les usages de l'IA dans notre société. Plusieurs mises en garde énergiques ont été faites à l'Europe pour éviter une position trop attentiste ou conservatrice par rapport à la Chine et aux États-Unis. L'Europe a choisi de se distinguer par une attention sociétale plus marquée et un meilleur respect de la vie privée. Cette approche est louable si elle est combinée avec une politique volontariste de

93 Un exemple intéressant d'initiative sur ce thème est le *Partnership on AI* lancé par un certain nombre de grands acteurs américains du domaine, dont les GAFAM : <https://www.partnershiponai.org/>

RECOMMANDATIONS

soutien dans les investissements qui sont nécessaires pour rattraper notre retard et avec le pragmatisme nécessaire au développement des écosystèmes mentionnés dans ce rapport.

Il faut pallier le déficit de communication des jeunes acteurs innovants suite à la faiblesse structurelle de la presse technique en France. Le nouveau modèle de diffusion de la presse en concurrence avec Internet a été défavorable à la presse technique en France, pour des raisons de taille critique et de nouveau modèle économique. Dans un écosystème d'innovation, l'accès à l'information est fondamental ; cette faiblesse de la presse technique complique l'accès à la notoriété pour les petits acteurs innovants. Dans le doute, les grands clients se tournent vers les technologies et fournisseurs américains, dont les méthodes de marketing sont plus massives et plus rodées. Il faut donc continuer à mettre en valeur les innovateurs, via des concours, des hackathons, des défis ou des événements, pour leur permettre d'émerger face à une « pensée unique » qui favorise naturellement les acteurs établis et les plus puissants médiatiquement.

Appendices

LISTE DES EXPERTS

- ▶ Patrick Albert — ex-ILOG (1/16)
- ▶ Gérard Sabah - ADT (2/16)
- ▶ Jean-Gabriel Ganascia — LIP6 (3/16)
- ▶ Yves Demazeau - AFIA (4/16)
- ▶ Norm Judah & Bernard Ourghanlian - Microsoft (5/16)
- ▶ Benoit Rottembourg - EuroDécision (6/16)
- ▶ Yann Lechelle - SNIPS (9/16)
- ▶ Pierre-Yves Oudeyer - INRIA (11/16)
- ▶ Jean-Louis Lezaun - IBM (12/16)
- ▶ Manuel Da Rocha - SOPRA (1/17)
- ▶ Cédric Sauviat - AFCIA (2/17)
- ▶ François Laburthe - Amadeus (3/17)
- ▶ Eunika Mercier — U. Lyon / EPITA (4/17)
- ▶ Francis Bach - INRIA (5/17)
- ▶ Jacques Serris - CGEJET (6/17)
- ▶ Michelle Sebag - LRI (7/17)
- ▶ Christophe Bourguignat — Zelros (10/17)
- ▶ Jean-Paul Smets — NEXEDI (12/17)
- ▶ Laurent Alexandre — ADT (3/18)

BIBLIOGRAPHIE

- [1] Académie des technologies. Big Data, un changement de paradigme peut en cacher un autre. Rapport, Décembre 2015.
- [2] Cathy O'Neil. *Weapons of Math Destruction : How Big Data Increases Inequality and Threatens Democracy*. Broadway Books, 2016.
- [3] Cecil Dijoux. *Ce que signifie l'avènement du numérique*. #hyperlean, 2017
- [4] David Barber. *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012.
- [5] Doc Searls. *The Intention Economy*. Harvard Business Review Press, 2012.
- [6] Eric Ries. *The Lean Startup*. Crown Business, 2011.
- [7] Eric Biernat, Michel Lutz. *Data Science : Fondamentaux et études de cas*. Eyrolles, 2015.
- [8] Erik Brynjolfsson, Andrew McAfee. *The Second Machine Age : Work, Progress and Prosperity in a Time of Brilliant Technologies*. Norton, 2014.
- [9] Foster Provost & Tom Fawcett. *Data Science for Business*. O'Reilly, 2013.

- [10] France IA. Rapport de synthèse. 2017. <http://www.enseignementsup-recherche.gouv.fr/cid114739/rapport-strategie-france-i.a.-pour-le-developpement-des-technologies-d-intelligence-artificielle.html>
- [11] Herbert Gintis. *Game Theory Evolving - A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, 2009.
- [12] Ian Goodfellow, Yoshua Bengio, Aaron Courville. *Deep Learning*. MIT Press, 2016.
- [13] INRIA. Artificial Intelligence — Current Challenges and Inria's engagement.in Livre blanc INRIA 2016
- [14] Ilarion Pavel, Jacques Serris. Modalité de régulation des algorithmes de traitement des contenus. Rapport du CGEJET, 2016.
- [15] Jean-Claude Heudin. *Comprendre le Deep Learning : Une introduction aux réseaux de neurones*. Scien e-book, 2016.
- [16] Jean-Gabriel Ganascia. *Le Mythe de La Singularité. Faut-il craindre l'intelligence artificielle ?* Le Seuil, 2017.
- [17] Jorgen W. Weibull. *Evolutionary Game Theory*. The MIT Press 1997
- [18] Marco Tinelli. *Le Marketing Synchronisé*. Eyrolles, 2012.
- [19] Max Tegmark. *Life 3.0 : Being Human in the Age of Artificial Intelligence*. Knopf, 2017.

BIBLIOGRAPHIE

- [20] Michio Kaku. *The Future Of The Mind*. Allen Lane, 2014.
- [21] Nassim Taleb. *Fooled by Randomness*. Random House, 2005.
- [22] OPECST. Pour une intelligence artificielle utile, maîtrisée et partagée. Mars 2017.
- [23] Pedro Domingos. *The Master Algorithm : How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, 2015
- [24] Ray Kurzweil. *How To Create a Mind — The secret of human thought revealed*. Penguin, 2013
- [25] Ray Kurzweil. *The Singularity is near — When Humans Transcend Biology*. Penguin, 2010.
- [26] S. Russel & P. Norvig, Eds. *Artificial Intelligence : A Modern Approach*. Prentice Hall Series. 2002.
- [27] Seth Stephens-Davidowitz. *Everybody Lies : Big Data, and What the Internet Can Tell Us About Who We Really Are*. Dey Street Books, 2017.
- [28] World Economic Forum. *Technology and Innovation for the Future of Production : Accelerating Value Creation*. 2017. http://www3.weforum.org/docs/WEF_White_Paper_Technology_Innovation_Future_of_Production_2017.pdf
- [29] White House. Preparing for the future of Artificial Intelligence Artificial Intelligence. Octobre 2016.

Glossaire

API	<i>(Application Programming Interface)</i> : Interface logicielle pour l'interconnexion applicative
ASIC	<i>(Application-Specific Integrated Circuit)</i> : processeur spécialisé pour certaines applications
Big Data	Approche de traitement des données informatique massives (mégadonnées)
CGEJET	Conseil général de l'économie, de l'industrie, de l'énergie et des technologies
Chatbots	Agents conversationnels
CNN	<i>(Convolutated Neural Networks)</i> Réseaux neuronaux convolutifs (voir 2.1.2.2)
Data mining	extraction de connaissance à partir des données
Deep Learning	Apprentissage automatique au moyen de réseaux neuronaux profonds
DNN	<i>(Deep Neural Nets)</i> Réseaux neuronaux profonds (nombre de couches)
GAFAM	Google, Apple, Facebook, Amazon, Microsoft
GDPR	<i>(General Data Protection Regulation)</i> : Règlement général

	sur la protection des données
GPU	(<i>Graphical Processing Unit</i>) : processeur spécialisé dans le traitement graphique <i>Hackathon</i> : Concours de programmation de courte durée
<i>Machine learning</i>	Apprentissage automatique
<i>Monte-Carlo</i>	Méthode probabiliste d'évaluation de problème
NLP	(<i>Natural Language Processing</i>) : Traitement du langage naturel
OPECST	Office parlementaire d'évaluation des choix scientifiques et techniques
Open Source	logiciel libre, dont le code source est accessible à tous (avec différentes formes de licences)
<i>Overfitting</i>	Surapprentissage ou sur-ajustement.
<i>Reinforcement Learning</i>	Apprentissage par renforcement (voir 2.1.2.3)
Rétropropagation	algorithmes des réseaux de neurones qui permettent de propager le signal attendu en sortie (classification) vers les neurones d'entrée.
RPA	(<i>Robotic Process Automation</i>) : Robots logiciels permettant l'automatisation des processus
SaaS	(<i>Software as a Service</i>) : Mise à disposition de services logiciels depuis le « cloud » (« nuage » de ressources informatiques accessibles par le réseau Internet).
<i>Spurious correlations</i>	Corrélations fortuites
<i>Supply Chain</i>	Gestion de la chaîne logistique
SVM	(<i>Support Vector Machine</i>) méthode d'apprentissage statistique (machine à support de vecteur), une généralisation des classifieurs linéaires.

Théorie des jeux évolutionnaire	une approche de calcul d'équilibres de Nash par itération qui s'appuie sur des méthodes évolutionnaires, semblables aux algorithmes génétiques, qui font évoluer une « population » de solutions approchées.
TPU	(<i>TensorFlow Processing Unit</i>) processeur spécialisé dans le calcul de réseau de neurones.
Transfer Learning	Apprentissage par transfert (entraîner le réseau sur une classe de problème pour en extraire des couches qui sont réutilisées pour une autre classe)

Table des matières

RÉSUMÉ	5
INTRODUCTION	9
ÉTAT DE L'ART	15
Comprendre l'Intelligence artificielle	15
<i>Les multiples formes de l'IA</i>	15
<i>Ce que l'IA sait faire aujourd'hui</i>	20
<i>Exemples d'application de l'IA</i>	28
Intelligence artificielle, apprentissage et mégadonnées	37
Pourquoi un rapport de l'Académie des technologies ?	40
<i>Les rapports récents</i>	40
<i>Ce qui fait consensus sur le domaine de l'Intelligence artificielle et de l'apprentissage</i>	43
QUELLE INTELLIGENCE ARTIFICIELLE POUR QUELS PROBLÈMES ?	47
Comprendre la boîte à outils	47
Savoir appliquer des méthodes simples à bon escient	49
Le processus d'extraction de valeur de la donnée	53

GARDER UN REGARD LUCIDE SUR LES PROGRÈS TECHNOLOGIQUES	57
Un vrai risque d'emballement qui pourrait conduire à un nouvel hiver	57
Le rôle clé de la pratique et de la technologie	59
Il reste des vraies barrières scientifiques	62
Il est difficile de prédire ce que l'on ne comprend pas	64
FAVORISER LE DÉVELOPPEMENT DE L'IA DANS LES ENTREPRISES	69
Comprendre les conditions de succès	69
Traquer les idées reçues	71
Construire une culture propre à chaque domaine d'activité	73
RECOMMANDATIONS	77
Soutenir l'effort de promotion engagé par « France IA »	77
Traiter les questions éthiques de l'IA le plus tôt possible	80
Promouvoir l'intelligence artificielle dans une perspective large	84
APPENDICES	88
Liste des experts	88
Bibliographie	89
Glossaire	92

Publications de l'Académie

(dix dernières années)

Avertissement

Les travaux de l'Académie des technologies sont l'objet de publications réparties en quatre collections¹ :

Les rapports de l'Académie : ce sont des textes rédigés par un groupe de l'Académie dans le cadre du programme décidé par l'Académie et suivi par le Comité des travaux. Ces textes sont soumis au Comité de la qualité, votés par l'Assemblée, puis rendus publics. On trouve dans la même collection les avis de l'Académie, également votés en Assemblée, et dont le conseil académique a décidé de la publication sous forme d'ouvrage papier. Cette collection est sous couverture bleue.

Les communications à l'Académie sont rédigées par un ou plusieurs Académiciens. Elles sont soumises au Comité de la qualité et débattues en Assemblée. Non soumises à son vote elles n'engagent pas l'Académie. Elles sont rendues publiques comme telles, sur décision du Conseil académique. Cette collection est publiée sous couverture rouge.

Les « Dix questions à... et dix questions sur... » : un auteur spécialiste d'un sujet est sélectionné par le Comité des travaux et propose dix à quinze pages au maximum, sous forme de réponses à dix questions qu'il a élaborées lui-même ou après discussion avec un journaliste de ses connaissances ou des collègues (Dix questions à...). Ce type de document peut aussi être rédigé sur un thème défini par l'Académie par un académicien ou un groupe d'académiciens (Dix questions sur...). Dans les deux cas ces textes sont écrits de manière à être accessibles à un public non-spécialisé. Cette collection est publiée sous une couverture verte.

¹ - Les ouvrages de l'Académie des technologies publiés entre 2008 et 2012 peuvent être commandés aux Éditions Le Manuscrit ([http : //www.manuscrit.com](http://www.manuscrit.com)). La plupart existent tant sous forme matérielle que sous forme électronique.
- Les titres publiés à partir de 2013 à 2016 sont disponibles en librairie et sous forme de ebook payant sur le site de EDP sciences ([http : //laboutique.edpsciences.fr/](http://laboutique.edpsciences.fr/)). À échéance de six mois ils sont téléchargeables directement et gratuitement sur le site de l'Académie.
- Les titres publiés depuis 2016 sont disponibles au secrétariat de l'Académie des technologies.
- Les publications plus anciennes n'ont pas fait l'objet d'une diffusion commerciale, elles sont consultables et téléchargeables sur le site public de l'Académie www.academie-technologies.fr, dans la rubrique « Publications ». De plus, l'Académie dispose encore pour certaines d'entre elles d'exemplaires imprimés.

PUBLICATIONS DE L'ACADÉMIE DES TECHNOLOGIES

Les grandes aventures technologiques françaises : témoignages d'un membre de l'Académie ayant contribué à l'histoire industrielle. Cette collection est publiée sous couverture jaune.

Par ailleurs, concernant les Avis, l'Académie des technologies est amenée, comme cela est spécifié dans ses missions, à remettre des Avis suite à la saisine d'une collectivité publique ou par auto saisine en réaction à l'actualité. Lorsqu'un avis ne fait pas l'objet d'une publication matérielle, il est, après accord de l'organisme demandeur, mis en ligne sur le site public de l'Académie.

Enfin, l'Académie participe aussi à des co-études avec ses partenaires, notamment les Académies des sciences, de médecine, d'agriculture, de pharmacie...

Tous les documents émis par l'Académie des technologies depuis sa création sont répertoriés sur le site www.academie-technologies.fr. La plupart sont peuvent être consultés sur ce site et ils sont pour beaucoup téléchargeables.

Dans la liste ci-dessous, les documents édités sous forme d'ouvrage imprimé commercialisé sont signalés par une astérisque. Les publications les plus récentes sont signalées sur le site des éditions. Toutes les publications existent aussi sous forme électronique au format pdf et pour les plus récentes au format ebook.

Liste des publications

AVIS DE L'ACADÉMIE

1. Évaluation de la recherche technologique publique – février 2013
2. L'usage de la langue anglaise dans l'enseignement supérieur – mai 2013
3. Les Académies d'agriculture, des sciences et des technologies demandent de restaurer la liberté de recherche sur les plantes génétiquement modifiées – mars 2014
4. La réglementation thermique 2012, la réglementation bâtiment responsable 2020 et le climat – novembre 2014
5. Les réseaux de chaleur – décembre 2014
6. Les enjeux stratégiques de la fabrication additive – juin 2015
7. Sur la loi relative à la "transition énergétique pour une croissance verte" – juin 2015
8. Les technologies et le changement climatique. Des solutions pour l'atténuation et l'adaptation – novembre 2015
9. Biodiversité et aménagement des territoires – décembre 2015
10. Aliments-santé. Implications pour l'industrie – mai 2016
11. Avis des Académies d'agriculture de France et des technologies sur la réglementation des mutagenèses ciblées en amélioration des plantes – juillet 2016
12. La détermination d'un prix de référence du carbone – janvier 2017
13. Modifier la réglementation thermique des bâtiments neufs afin de baisser les émissions de gaz à effet de serre au moindre coût – juillet 2017
14. Une stratégie robotique pour réindustrialiser la France – juillet 2017
15. Le passage à l'échelle industrielle de la production de cellules souches à usage thérapeutique - Avis commun de l'Académie des technologies et de l'Académie nationale de médecine- janvier 2018

PUBLICATIONS DE L'ACADÉMIE DES TECHNOLOGIES

RAPPORTS DE L'ACADÉMIE

1. *Le véhicule du futur – septembre 2012 (publication juin 2013)
2. *Freight systems (version anglaise du rapport 10 le transport de marchandises) – novembre 2012
3. *Energy vectors – novembre 2012 (version anglaise du numéro 13)
4. *Energy Efficiency in Buildings and Housing – novembre 2012 (version anglaise du numéro 11)
5. *Les grands systèmes socio-techniques / Large Socio-Technical Systems – ouvrage bilingue , juillet 2013
6. *Première contribution de l'Académie des technologies au débat national sur l'énergie / First contribution of the national academy of technologies of France to the national debate on the Future of energies supply – ouvrage bilingue , juillet 2013
7. Renaissance de l'industrie : construire des écosystèmes compétitifs fondés sur la confiance et favorisant l'innovation - juillet 2014
8. Le Méthane : d'où vient-il et quel est son impact sur le climat ? – novembre 2014
9. Impact des TIC sur la consommation d'Énergie à travers le monde – 2015
10. Big data : un changement de paradigme peut en cacher un autre – décembre 2015
11. Le biogaz – mars 2016
12. Les technologies du changement climatique : des solutions pour l'atténuation et l'adaptation – avril 2016
13. L'enseignement professionnel – mai 2016
14. La perception des risque – juin 2016
15. Les aliments santé : avancées scientifiques et implications industrielles – octobre 2016
16. Quel prix de référence du CO2 ? – mars 2017
17. Les technologies du changement climatique – des solutions pour l'atténuation et l'adaptation (addendum) – avril 2017
18. Innovation ouverte et PME – juin 2017
19. Innovation – Croissance – Emploi . Rapport du séminaire 2016 de l'Académie des technologies - septembre 2017
20. Approche des processus fondamentaux de l'apprentissage – septembre 2017
21. Industrie du futur : du système technique 4.0 au système social – décembre 2017
22. Recommandations pour développer et valoriser la formation professionnelle initiale et continue – décembre 2017
23. Technologies et territoires d'innovation : quatre réflexions pertinentes et impertinentes - février 2018
24. Renouveau de l'intelligence artificielle et de l'apprentissage automatique, mars 2018

COMMUNICATIONS À L'ACADÉMIE

1. Renaissance de l'industrie : recueil d'analyses spécifiques - juillet 2014
2. Réflexions sur la robotique militaire – février 2015
3. Le rôle de la technologie et de la pratique dans l'enseignement de l'informatique - novembre 2015
4. Le pétrole , quelles réserves , quelles productions et à quels prix ? - janvier 2017
5. L'incidence sur la santé humaine des différentes sources de production d'énergie électrique: évaluation sur les cinquante dernières années - juin 2017
6. Technologies et territoires d'innovation - février 2018

PUBLICATIONS DE L'ACADÉMIE DES TECHNOLOGIES

DIX QUESTIONS POSÉES À...

1. Les OGM - 10 questions posées à Bernard Le Buanec - avril 2014
2. *Comment bien se nourrir en respectant la planète et notre santé ? - 10 questions posées à Pierre Feillet - juin 2016

GRANDES AVENTURES TECHNOLOGIQUES

1. *Le Rilsan – par Pierre Castillon – octobre 2006
2. *Un siècle d'énergie nucléaire – par Michel Hug – novembre 2009

HORS COLLECTION

1. Quel avenir pour les biocarburants aéronautiques ? – Rapport commun avec l'Académie de l'Air et de l'Espace – juillet 2015
2. La mise en œuvre en France des techniques de séquençage de nouvelle génération . Rapport commun avec l'Académie de médecine – février 2016
3. Le passage à l'échelle industrielle de la production de cellules souches humaines à usage thérapeutique - Rapport commun avec l'Académie nationale de médecine - janvier 2018