



Pirmin Lemberger - **Data Scientist & Directeur DataLab Weave**
Jonathan Lepan - **Senior Manager**
Olivier Reisse - **Fondateur de weave Business Technology**

LIVRE BLANC

IA

OÙ EN SOMMES NOUS ?

INTELLIGENCE ARTIFICIELLE, OÙ EN SOMMES NOUS ?

Pirmin Lemberger
Data Scientist & Directeur DataLab Weave

Jonathan Lepan
Senior Manager

Olivier Reisse
Fondateur de weave Business Technology

Un livre blanc
by Weave Business Technology



37 rue du rocher, 75008 Paris
+33 (0)1 58 44 58 44
weave.eu

A network diagram background consisting of numerous grey circular nodes of varying sizes connected by thin, light grey lines. The nodes are scattered across the page, with a higher density in the upper left and lower right areas. The lines create a complex web of connections between the nodes.

BIOGRAPHIE DES AUTEURS



Pirmin Lemberger

Pirmin est Data scientist, physicien de formation, il anime et dirige le data lab de *Weave*. Ses tribunes paraissent dans *IT for Business* et le *JDN* ainsi que sur *Ristretto* le blog du cabinet *Weave*. Il est l'auteur principal du livre *Big Data et Machine Learning* paru chez *Dunod* en 2016.

pirmin.lemberger@weave.eu



Jonathan Lepad

Jonathan est Senior Manager au sein de l'entité *Business Technology* du cabinet *Weave*, Ingénieur avec une spécialisation en Ingénierie de la Connaissance, il intervient sur les nouvelles pratiques digitales dans le mode de fonctionnement des entreprises.

jonathan.lepad@weave.eu



Olivier Reisse

Olivier membre du comité exécutif de *weave* est associé fondateur de la practice *Business Technology*. Depuis 15 ans au sein de *Weave*, Olivier travaille à développer une vision stratégique de la place et du rôle de la technologie dans l'entreprise. Nous les accompagnons sur la voie des changements culturels, organisationnels et managériaux induits par l'évolution des usages numériques.

olivier.reisse@weave.eu



SOMMAIRE

1. L'IA en quête de définition	
1.1 Pourquoi une définition de l'IA est difficile	10
1.2 Quels objectifs pour l'IA?	13
1.3 Comment caractériser une IA?	15
2. Progrès et lacunes de l'IA	
2.1 Une histoire mouvementée	21
2.2 Les percées en attente d'un Einstein de l'IA	26
2.3 La compréhension du langage naturel	29
3. Les usages	
3.1 Aujourd'hui	33
3.2 Bientôt	37
3.3 Demain?	41
4. Une éthique et un droit pour l'IA	
4.1 Comment répartir équitablement les richesses créées par l'IA?	43
4.2 Comment garantir la neutralité et la légalité d'un système prédictif?	45
4.3 Quelles responsabilités en cas d'accident?	47
4.4 Une nouvelle forme d'assistanat?	48
4.5 Et si l'homme se rapprochait des machines (plutôt que l'inverse)?	49
5. L'IA et le métier du conseil	
5.1 Avant-projet	52
5.2 Projet	54
Conclusion	56
Travaux cités	58



L'IA EN QUÊTE DE DÉFINITION

À condition de pas s'embarasser de trop de scrupules sémantiques, la définition de l'intelligence artificielle (IA) est aisée : c'est l'ensemble des disciplines techniques et scientifiques qui permettent de reproduire certains processus cognitifs humains comme l'apprentissage, l'intuition, l'auto-amélioration, la créativité, la planification de tâches ou encore la compréhension du langage naturel. Les plus optimistes y ajouteront pour leur part, en les plaçant dans un futur indéfini, des attributs propres à la conscience comme la volonté et les émotions.

Face à l'effroi métaphysique qui peut saisir l'esprit devant cette ambition prométhéenne, nous tâcherons de garder la tête froide. En particulier nous distinguerons ce qui existe aujourd'hui de qui reste à réaliser, ce qui relève de la spéculation de ce qui est avéré scientifiquement. Nous éviterons également de nous livrer à toute forme de **prophétie**, laissant cette activité aux prophètes du transhumanisme qui y excellent.

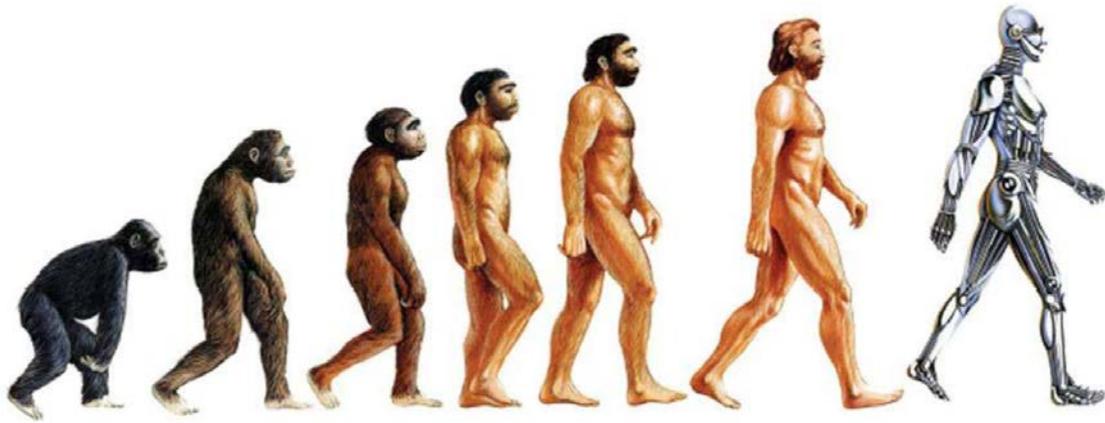


Figure 1 : L'évolution de l'humanité selon les transhumanistes

Ces précautions prises, il faut cependant bien reconnaître que, si d'aventure, le projet de construire une IA de niveau humain devait un jour aboutir, il ne s'agirait ni plus ni moins de l'événement le plus significatif de l'histoire de l'humanité. Voilà au moins une raison suffisante pour y réfléchir de manière lucide et informée dès maintenant. Quelle que soit l'aboutissement de cette aventure gageons que nous en apprendrons autant sur ce qui fait l'essence de l'humanité que sur les capacités cognitives d'une machine.

Une définition plus rigoureuse de l'IA que celle mentionnée en introduction se heurte néanmoins à plusieurs difficultés, voyons pourquoi.

1.1 POURQUOI UNE DÉFINITION DE L'IA EST DIFFICILE

La pluridisciplinarité

Les principales disciplines qui ont contribué à l'émergence de l'IA au cours des soixante dernières années ont chacune des questions, des méthodes et des objectifs qui leurs sont propres. En voici une liste non exhaustive :

- **La philosophie** pose par exemple la question de l'articulation entre l'esprit et la matière ou celle de savoir comment l'on passe de la connaissance à l'action.
- **Les mathématiques** s'intéressent aux règles formelles qui permettent de garantir la validité d'un raisonnement (qui est une forme spécifique d'intelligence). Sur un plan conceptuel l'informatique est d'ailleurs née avec les travaux d'Alan Turing qui était en quête d'une définition rigoureuse de ce qu'est un calcul. La théorie des probabilités permet quant à elle de savoir comment agir rationnellement dans les situations incertaines dans lesquelles sont amenées à évoluer des agents intelligents.
- **L'économie** a pour objectif, entre autres, de comprendre les performances collectives d'un ensemble d'agents rationnels qui cherchent chacun à optimiser leurs propres revenus ou leur bien-être. L'une des découvertes surprenantes faite dans le cadre de la théorie des jeux est qu'un agent rationnel aura parfois intérêt à adopter une stratégie aléatoire. Savoir prendre une décision lorsque les bénéfices sont différés dans le temps ou lorsque l'on autorise des décisions qui ne sont pas optimales font partie des questions qui restent ouvertes à ce jour.
- **Les neurosciences** étudient le fonctionnement du cerveau et comment les différentes fonctions cognitives sont reliées à l'interconnexion de certains groupes de neurones. Le cortex visuel des mammifères a inspiré certaines architectures de systèmes de neurones artificiels utilisés pour la reconnaissance de formes.
- **La cybernétique** est l'étude des mécanismes autorégulés, une propriété qui était jusque-là l'apanage des seuls organismes vivants. La version moderne de la théorie du contrôle consiste à construire des systèmes qui optimisent la valeur d'une certaine fonction de coût au cours du temps.

L'IA s'est construite durant un demi-siècle en émancipation de toutes ces disciplines en définissant progressivement ses propres objectifs comme par exemple celui de concevoir des machines capables d'évoluer de manière

autonome dans des environnements changeants.

Une histoire chaotique

L'histoire de l'IA est jalonnée d'une succession d'ambitions grandioses et d'espoirs déçus, nous y reviendrons dans la section 2.1. Dans les années 1950 les chefs de files de l'IA naissante s'imaginaient par exemple que des progrès significatifs sur la conception de systèmes qui pensent comme les humains pourraient être réalisés durant un séminaire de deux mois ([voir page 21](#)). Des approches techniques et conceptuelles disparates ont ainsi pendant longtemps coexisté et empêché de concevoir l'IA comme une discipline unifiée.

Une absence d'unanimité

Même si le sujet est dorénavant mieux structuré comme nous le verrons dans la section 1.2, il n'existe à ce jour aucune unanimité parmi les spécialistes sur les objectifs que l'on devrait assigner à l'IA. Nous décrivons ci-dessous ([voir page 13](#)) quatre points de vue développés par Russel et Norvig dans leur ouvrage monumental *Artificial Intelligence a Modern Approach* (Norvig, et al., 2009). A l'origine de cette absence d'unanimité se trouve naturellement la difficulté à définir ce qu'est l'intelligence ou même seulement certains de ses aspects comme la créativité, l'intuition ou la capacité d'abstraction.

L'IA c'est toujours le futur

Une autre difficulté que le rencontre lorsqu'on cherche à définir ce qu'est une IA est le caractère subjectif d'une telle définition. Les problèmes dont on estime qu'ils exigent une «authentique» IA pour être résolus varient en effet au cours de l'histoire. Ainsi une machine capable de battre un grand maître aux échecs aurait été considérée il y a cinquante ans comme «intelligente». Une telle machine fut effectivement construite en 1997 lorsque *Deep Blue* démontra sa supériorité dans un match contre le champion Kasparov. Du statut de prouesse on passa alors très rapidement à celui de produit grand public et plus personne ne songerait désormais qu'un programme d'échecs exécuté sur un PC soit doté d'une «authentique» intelligence.

Il y a encore deux ans peu de gens auraient parié sur la possibilité qu'un ordinateur puisse battre un grand maître au jeu de Go, du moins avant une bonne vingtaine d'années. L'avis des experts était qu'une réelle maîtrise de ce jeu requérait une intuition, une imagination (voire pour certains une sensibilité

mystique) inaccessibles à une machine. C'est pourtant chose faite depuis le début de l'année (2016) avec le système *AlphaGo* de *Google* qui a battu le champion Lee Sedol lors d'un tournoi historique. Une nouvelle fois cette performance n'est désormais plus considérée comme de l'IA.

On pourrait aisément multiplier les exemples. Bref, l'IA est souvent perçue comme ce qui n'a pas encore été réalisé. Il y a donc fort à parier qu'il n'y aura jamais de moment Eurêka dans la quête de l'IA mais plus vraisemblablement une succession de phases de progrès et d'accoutumance qui s'étaleront sur plusieurs décennies.

Illusion ou réalité ?

Comment reconnaître une intelligence généraliste (on parle parfois d'**IA forte**) d'un simple simulacre qui agit de manière à faire illusion ? L'idée qu'un simulacre d'IA serait immanquablement détecté après une mise à l'épreuve suffisamment longue est l'essence même du célèbre **test de Turing**. Celui-ci consiste à mettre à l'épreuve un système prétendument intelligent pour mesurer sa capacité à simuler une conversation humaine sans contact physique. Bien entendu aucun système à ce jour n'a encore passé ce test. Toutefois certaines applications pourraient dès aujourd'hui instiller un doute quant à leur capacité à « comprendre le monde ». Un système comme le logiciel de *Google* capable de décrire par un texte la scène représentée par une photo (voir page 34) est-il intelligent dans le sens où nous le souhaiterions ? La réponse est : probablement non. Nous vivons cependant une période passionnante dans la mesure où la réponse à cette question n'est désormais plus tout à fait évidente.

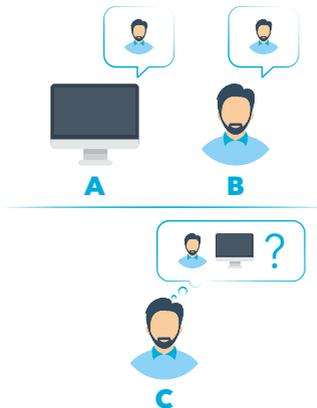


Figure 2 : Le schéma du test de Turing

1.2 QUELS OBJECTIFS POUR L'IA ?

Face aux difficultés à définir clairement ce qu'est un IA, une possibilité consiste à mieux définir les objectifs que l'on attend d'un tel système. En d'autres termes, il faut poser la question de savoir quels problèmes une IA devra être capable de résoudre. On peut schématiquement regrouper les approches de l'IA en quatre catégories représentées dans le tableau ci-dessous. Celui-ci définit deux axes : l'axe vertical distingue « penser » et « agir » et l'axe horizontal distingue l'idéal de comportement visé : un comportement « humain » ou un comportement rationnel (Norvig, et al., 2009).

PENSER COMME UN HUMAIN

Il s'agit de créer des « machines » dont le mode de fonctionnement interne reproduit le mode de pensée humain.

PENSER RATIONNELLEMENT

Il s'agit de créer des machines dont le mode de fonctionnement interne se base sur des déductions logiques.

AGIR COMME UN HUMAIN

Il s'agit de créer des machines capables de reproduire un comportement qui s'apparente, vu de l'extérieur, à celui d'un humain.

AGIR RATIONNELLEMENT

Il s'agit de construire des agents dont le comportement, vu de l'extérieur, apparaît rationnel.

Penser comme un humain

Dans cette approche on cherche à comprendre comment fonctionne l'esprit humain pour le modéliser, l'espoir étant de pouvoir, un jour peut-être, reproduire ses facultés cognitives dans un système artificiel. Les **sciences cognitives** suivent cette voie et élaborent des méthodes expérimentales sur des humains ou des animaux pour développer un modèle de l'esprit susceptible d'être testé expérimentalement. Certaines approches précoces de l'IA, comme celle du *General Problem Solver* élaborée au début des années 1960 par Herbert Simon, ont ainsi essayé de reproduire les étapes d'un raisonnement humain.

Agir comme un humain

Le **test de Turing** que nous avons déjà évoqué propose une définition opérationnelle de l'intelligence sans se préoccuper des processus internes à

la machine. Construire un système capable de passer un tel test impliquerait cependant de résoudre une cascade de problèmes parmi les plus ardues de l'IA: le traitement du langage humain, la représentation des connaissances, le raisonnement et l'apprentissage automatique. A ce jour la recherche en IA privilégie plutôt la compréhension des mécanismes qui sous-tendent l'intelligence sur la reproduction d'une intelligence capable de passer le test de Turing. La démarche est en l'occurrence similaire à l'aéronautique dont les succès ont reposé sur une compréhension approfondie des lois de l'aérodynamique plutôt que sur l'imitation du vol des oiseaux.

Penser rationnellement

La pensée rationnelle repose avant tout sur la logique. Son histoire débute avec Aristote et aboutit à la logique mathématique moderne. Alors que l'arithmétique traite de la manipulation des nombres, la logique mathématique formalise les enchaînements de propositions de portées plus ou moins générale. **L'approche logiciste** de l'IA part du postulat que toutes les opérations mentales se réduisent en dernière analyse à des opérations logiques. Cette approche se heurte toutefois à des deux difficultés principales.

- D'une part il est difficile de formaliser logiquement un problème mal défini ou même un problème dont les données sont entachées d'incertitudes.
- D'autre part, il existe un gouffre entre résoudre un problème en principe, c.à.d. avec des ressources de calcul à priori infinies, et le résoudre en pratique, c.à.d. en utilisant des ressources limitées durant un laps de temps acceptable afin d'obtenir une solution acceptable plutôt qu'une solution parfaite qui résulterait d'un raisonnement strictement logique.

Agir rationnellement

Une action rationnelle peut être une conséquence d'un raisonnement rationnel... ou non ! Dans certaines situations d'urgence par exemple, une action rationnelle n'est pas nécessairement le résultat d'un long processus de délibération mais résulte plutôt d'un réflexe appris par expérience. Dans d'autres situations encore, une action est requise alors qu'aucun raisonnement rationnel n'existe qui permettrait de la déduire. L'approche par action rationnelle est donc plus générale que l'approche logiciste. Elle présente par ailleurs l'avantage de permettre une définition plus scientifique que les approches qui visent à agir ou à penser comme un humain, difficilement susceptibles de définitions rigoureuses. C'est le point de vue moderne de l'IA développé dans *Norvig, et al., 2009* :

«L'IA est la science de la conception d'agents rationnels, des agents qui optimisent la valeur d'attente d'une certaine notion d'utilité en fonction des perceptions passées de leur environnement.»

L'action rationnelle n'est en réalité qu'une première étape vers l'objectif plus ambitieux et plus réaliste qui serait d'agir dans des environnements complexes avec une **rationalité limitée**¹ sous contrainte de délais et de ressources, ce que les humains ont à faire tout au long de leur existence.

1.3 COMMENT CARACTÉRISER UNE IA ?

La notion d'intelligence n'étant pas bien définie pour une IA, il est illusoire de vouloir la mesurer au moyen d'une métrique qui serait conçue comme une sorte de «QI adapté aux machines». On peut cependant décrire plusieurs axes qui permettent de caractériser une IA (*Loukides, et al., 2016*).

La profondeur

Considérons une IA spécialisée comme *AlphaGo*. Sa profondeur pourrait être définie par exemple comme la probabilité de remporter un tournoi contre un grand maître. Il s'agit par conséquent d'une mesure de performance spécifique au champ d'application pour lequel elle a été conçue. On pourrait aussi évaluer une telle IA spécialisée par les temps de résolution de quelques problèmes étalons. Il n'est pas question en revanche de comparer les performances du système *Deep Blue* avec celles d'*AlphaGo*.

Le mode d'apprentissage

Le machine learning (ML) est une composante essentielle de beaucoup de systèmes intelligents dont on attend qu'ils soient capables d'apprendre tout ou partie de leur comportement à partir d'exemples. Voici les trois principaux modes d'apprentissage définis par le machine learning :

- **L'apprentissage supervisé** regroupe tous les algorithmes capables d'apprendre une relation entre certaines variables prédictives (des images d'animaux pour fixer les idées) et une caractéristique que l'on souhaite prédire (le nom de l'animal représenté sur l'image p.ex.) à partir d'une liste d'exemples correctement identifiés ou mesurés. C'est aujourd'hui la partie du ML la plus développée en pratique et la mieux comprise sur le plan théorique. Les algorithmes les

¹ Herbert Simon a reçu en 1982 le prix Nobel d'économie pour sa théorie descriptive de la rationalité limitée.

plus performants pour apprendre des relations complexes sont aujourd'hui les réseaux de neurones (RN) profonds (deep learning). Ils semblent capables, dans une certaine mesure, de découvrir par eux-mêmes certains concepts abstraits dans données complexes comme des images. Ainsi un algorithme de reconnaissance d'un animal découvrira implicitement le concept « nombre de pattes ».

- **L'apprentissage non-supervisé** regroupe les algorithmes du ML capables de découvrir des structures dans un ensemble de données. Contrairement au ML supervisé il n'existe pas de formulation précise de ce problème si on le considère dans toute sa généralité. Il existe cependant un grand nombre d'algorithmes dit de **clustering** capables de regrouper entre elles des observations similaires. Ces techniques permettent par exemple d'identifier des niches d'un marché en regroupant des clients par catégories selon leurs profils. Les algorithmes de deep learning semblent capables, dans une certaine mesure, de découvrir des concepts abstraits dans des images. Ainsi un algorithme de reconnaissance d'un animal découvrira implicitement le concept « nombre de pattes ».

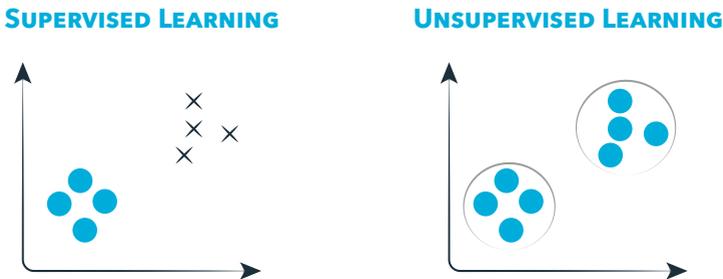


Figure 3 : Schéma des modes d'apprentissage supervisé et non-supervisé

- **L'apprentissage par renforcement** se situe en quelque sorte à mi-chemin entre les deux modes précédents. On n'alimente pas l'algorithme avec des exemples correctement classés mais on cherche à construire des systèmes qui évoluent dans un environnement complexe en maximisant un indicateur que l'on peut envisager comme une forme de récompense (ou un indice de confort si l'on préfère). Contrairement au ML supervisé le feedback est cette fois indirect. Il incite le système à trouver un bon compromis entre l'exploration de son environnement et l'exploitation de l'information qu'il a déjà acquise.

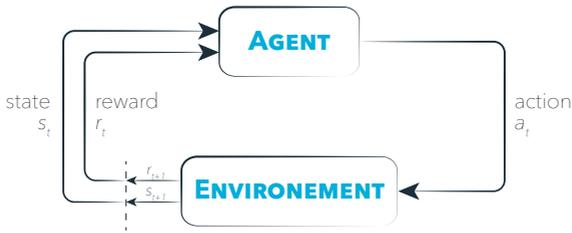


Figure 4 : Le schéma de rétroaction d'un environnement sur un agent qui définit l'apprentissage par renforcement.

L'immense majorité des applications actuelles de l'IA relève de l'apprentissage supervisé, qu'il s'agisse de systèmes de recommandation, de traduction automatique, de reconnaissance d'image ou de la parole.

L'étendue

Un système expert ou une **IA étroite** est spécialisé dans la résolution d'une catégorie bien précise de problèmes. Toutes les IA créés à ce jour sont de ce type. Il y a fort à parier qu'une IA généraliste (strong IA) ne s'obtiendra pas par simple juxtaposition d'une vaste collection d'IA étroites mais que de nouveaux principes restent à découvrir .

Quelques pas modestes ont cependant été fait en direction d'une IA généraliste en combinant des techniques d'apprentissage par renforcement avec celles du deep learning. Ainsi l'article *Volodymyr Mnih, 2015* décrit comment une équipe de *DeepMind* est parvenue à réaliser récemment un système capable d'apprendre à jouer à une cinquantaine de jeux vidéo Atari à une niveau super-humain et ceci en n'exploitant que les informations visuelles (sous forme de pixels) fournies par l'écran du jeu. En étant résolument optimiste on peut y voir, peut-être, les prémises d'une intelligence un peu moins étroite.

L'autonomie

Une IA peut assister un humain dans une tâche, à l'instar d'un GPS guidant un conducteur, sans pour autant être autonome face à toute la complexité de son environnement. Le système de pilotage d'une voiture sans conducteur est en revanche un système autonome qui n'exigera, en principe, aucune intervention humaine.

La présence de conscience

La question ne se pose pas, bien évidemment, avec les systèmes actuels ni, vraisemblablement, avec ceux qui seront conçus dans un avenir proche. Aussi nous aborderons ce point de manière très superficielle.

Il n'existe pas à l'heure actuelle de théorie de la conscience et personne ne comprend d'un point de vue scientifique la nature exacte de ce phénomène si bien que le sujet reste largement spéculatif. Nous prenons le terme ici dans l'acception commune d'une « présence objective d'une subjectivité » capable de ressentir, de se poser des questions et de vouloir. Sans entrer dans la question délicate de la détection d'une conscience, on peut néanmoins se demander si elle pourrait être utile et, si oui, dans quelles circonstances. D'innombrables applications de l'IA n'ont à l'évidence aucune utilité d'une conscience et pourrait même, assez ironiquement, bénéficier de la garantie d'en être dépourvu ! En effet une l'IA n'a pas à méditer sur sa propre existence mais à résoudre efficacement les problèmes qui lui sont assignés. En revanche, pour d'autres applications plus ambitieuses qui exigeraient par exemple de comprendre le sens d'un message ou d'interpréter un contexte physique ou social de la vie courante, la question reste ouverte.

Insistons pour conclure ce premier chapitre que l'intelligence que nous attribuons intuitivement à une IA dépourvue de conscience, n'est bien évidemment **qu'une projection de notre conscience humaine** sur le comportement d'un système qui, en réalité, ne procède qu'à un traitement élaboré de l'information.



PROGRÈS ET LACUNES DE L'IA

2.1 UNE HISTOIRE MOUVEMENTÉE

Voici quelques jalons de l'histoire de l'IA qui permettront de mieux situer les efforts de la recherche contemporaine.

L'enthousiasme des pionniers - avant 1966

En 1936 Allan Turing cherche à formaliser ce qu'est un calcul². Il introduit à cette occasion un concept révolutionnaire en mathématiques, appelé par la suite la **Machine de Turing**, conçu comme un calculateur universel. C'est une première tentative, sur un plan conceptuel, d'automatiser certains processus mentaux, ceux du calcul en l'occurrence. Un autre procédé de calcul universel, équivalent, basé sur des **réseaux de neurones** (RN) artificiels, a été étudié dès les années 1950. La création de la discipline de l'IA est traditionnellement associée à une conférence qui s'est tenue en 1956 à Dartmouth. Son objectif affiché illustre bien l'enthousiasme, rétrospectivement excessif, des débuts de l'IA :

« Nous nous proposons de réunir durant deux mois une dizaine d'experts pour essayer de concevoir des machines apprenantes qui utilisent le langage, forment des abstractions et résolvent des problèmes d'ordinaire réservés aux humains. Le postulat de base sera qu'il est possible de décrire de manière si précise tous les aspects de l'intelligence et de l'apprentissage qu'on pourra les simuler. »

Inutile de préciser que les percées escomptées ne furent pas réalisées. Un grand nombre de chercheurs en IA ont cependant fait connaissance à cette occasion.

En 1958, J. McCarthy inventa *LISP* un nouveau type de langage de programmation qui opère sur des expressions symboliques plutôt que sur des nombres. LISP restera durant 30 ans le langage de prédilection des chercheurs en IA qui adoptèrent l'**approche** dite **logiciste** à l'IA. Celle-ci part du postulat que l'on peut représenter tout raisonnement en logique formelle.

Le même année McCarthy publia un article historique *Programs with Common Sense* dans lequel il décrit un système hypothétique qui peut être vu comme un système d'IA complet et qui reste en grande partie d'actualité. Réduire au maximum la complexité d'un problème pour en saisir l'essence est une stratégie classique dans toute recherche scientifique. L'approche par **micro-mondes** a été élaborée dans cet esprit, elle consiste à définir un univers simplifié à l'extrême, constitué de formes géométriques élémentaires destinées à être manipulées par

² Il le fait dans la foulée d'un des problèmes proposés par Hilbert qui est le problème de la décision : existe-t-il une méthode systématique pour déterminer si une certaine proposition est démontrable ou non dans un système d'axiomes donné ?

un agent « intelligent ». Cet agent doit comprendre cet environnement ainsi que les ordres qu'on lui transmet. Il doit être en mesure de planifier puis d'exécuter des actions qui correspondent à l'injonction fournie. Des problèmes comme la vision, la planification de tâches ou la compréhension du langage naturel ont été abordés dans ce cadre expérimental.

Le **perceptron** que l'on peut envisager comme un neurone artificiel fut inventé par F. Rosenblatt en 1957. C'est un algorithme élémentaire de classification binaire capable d'apprendre à partir par expérience.

ELIZA développé au *MIT* dans les années 1964-66 est le premier chatbot de l'histoire, il tentait de simuler une conversation avec un psychanalyste.

Le temps des désillusions 1966 - 1986

L'espoir né avec les micro-mondes fut toutefois de courte durée car, hélas, cette approche ne concrétisa pas le premier pas tant espéré vers la construction de machines intelligentes. Ainsi le logiciel **SHDRLU**, développé au *MIT* par Terry Winograd, qui constitua la première tentative de compréhension du langage naturel dans ce cadre restreint, s'avéra être un échec. En effet, le passage à l'échelle vers des situations plus réalistes que celles des micro-mondes s'accompagnait systématiquement d'une **explosion combinatoire** qu'aucun progrès matériel pourrait jamais prendre en charge. Le constat amer était qu'en IA il existe un monde entre résoudre en problème « en principe » et le résoudre « en pratique ».

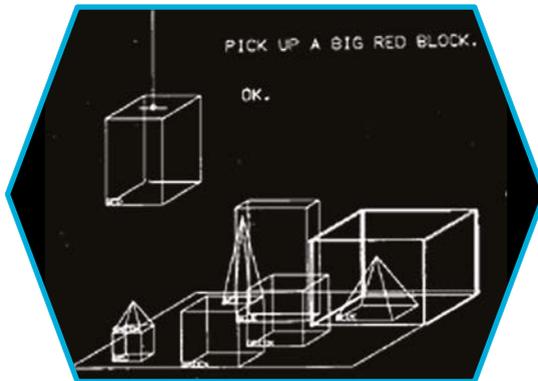


Figure 5 : Le micro-monde manipulé par le logiciel SHDRLU

Dans un cadre plus général que celui des micro-mondes les tentatives de

³ On cite souvent l'exemple classique de traduction bancale du français vers le russe de « L'esprit est fort mais la chair est faible » par « La vodka est bonne mais la viande est pourrie ».

compréhensions du langage achoppèrent face aux **ambiguïtés sémantiques** qu'aucun traitement purement syntaxique n'était capable de résoudre³. En effet une traduction correcte d'un texte nécessite, même lorsqu'elle est restreinte à un domaine spécifique, une forme de culture générale ou de bon sens. Il devint dès lors évident qu'une authentique compréhension du langage exigerait une connaissance générale du monde sur un plan physique et social.

En 1969 M.L. Minsky et S. Papert (Papert, 1972) publièrent un ouvrage qui pointait du doigt certaines limitations théoriques du perceptron. Ces limitations furent hélas indument interprétées comme s'appliquant par extension à tous les RN ce qui porta un coup fatal aux investissements dans la recherche dans cette discipline. Paradoxalement, au même moment, un algorithme clé pour l'entraînement des RN fut découvert : l'algorithme de **rétro-propagation**. Il fut par la suite redécouvert à plusieurs reprises durant les années 1980 et contribua à la résurgence actuelle des RN.

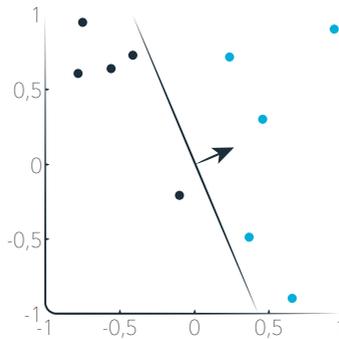


Figure 6 : L'algorithme du perceptron permet de trouver par apprentissage itératif un (hyper-)plan séparant deux groupes de données

Durant cette période furent également conçus les premiers systèmes experts basés sur des collections de règles. Un système de diagnostic médical comme MYCIN avec ses 450 règles s'est montré aussi performant que certains experts. Selon les systèmes, les règles utilisées étaient soit déduites à partir de connaissances fondamentales (comme la chimie p.ex.), soit recueillies laborieusement auprès d'experts humains (avec une longue expérience dans le diagnostic médical p.ex.).

Une approche plus scientifique 1986 - présent

L'approche de l'AI par les RN est qualifiée d'approche **connexionniste**. Elle se confronte à l'approche **logiciste** déjà évoquée et à l'approche **symbolique** qui permet l'élaboration de systèmes experts qui exploitent des collections de règles intelligibles par des humains. Les approches connexionnistes et symboliques sont aujourd'hui considérées comme complémentaires pour fournir un modèle de la connaissance.

Après les années d'expérimentations des pionniers, souvent basées sur des intuitions et des essais de multiples idées, cette période inaugura un retour à une démarche scientifique plus classique, basée désormais sur des résultats mathématiques rigoureux, des résultats expérimentaux quantifiables et des analyses statistiques détaillées. En 1997 *Deep Blue* bat Kasparov au jeu d'échec en utilisant essentiellement une méthode d'exploration arborescente des positions.

Un focus sur les données 2001 - présent

Avec l'avènement d'**Internet** les technologies d'IA ont été utilisées dans un premier temps par les moteurs de recherche et par les systèmes de recommandation devenus indispensables pour personnaliser l'expérience utilisateur de clients à qui l'on offre désormais des milliers voire des millions d'articles.

Internet a par ailleurs changé la donne en rendant accessibles d'énormes quantité de données, qu'il s'agisse de corpus de textes pour entraîner des systèmes de traduction automatique ou de collections d'images pour entraîner des algorithmes de reconnaissance de formes.

« Un fait marquant de cette période a été le constat qu'il souvent préférable d'augmenter le volume d'un jeu de données plutôt que de peaufiner les algorithmes⁴. »

Cette démarche par «force brutale», qui mise avant tout sur la quantité de données, atteindra peut-être un jour ses limites mais elle a permis de résoudre partiellement le problème de **désambiguïsation** lexicale mentionné précédemment et ceci sans qu'il soit nécessaire de construire explicitement de longues listes d'exemples et de règles.

Des algorithmes de complétion automatique d'images ont fait apparaître des effets de seuil surprenants, passant d'une performance médiocre à un niveau excellent, lorsque le nombre d'images utilisées pour entraîner l'algorithme

⁴ Un dicton résume bien cet état de fait : « Better data outweighs clever maths ! »

passait de quelques milliers à plusieurs millions.

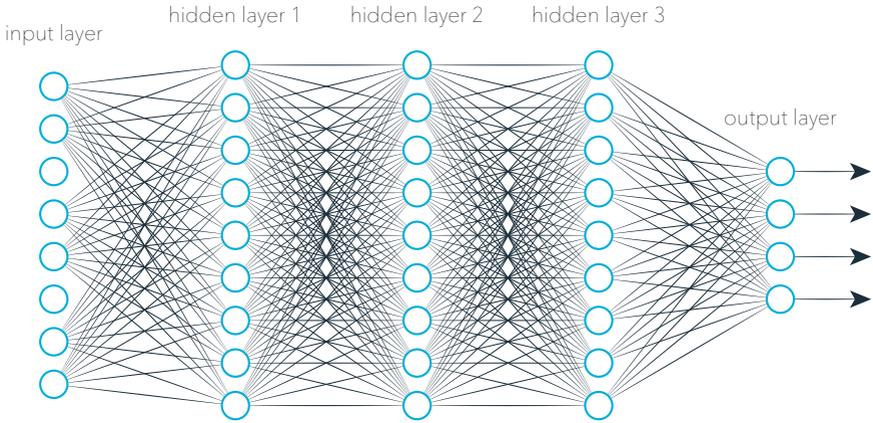


Figure 7 : Le deep learning exploite des réseaux de neurones multicouches que l'on sait désormais entraîner de manière efficace.

En 2011, le système *Watson* d'IBM bat la meilleure équipe humaine au jeu de *Jeopardy*, un jeu qui exige une grande culture générale et une aptitude significative à la désambiguïsation lexicale.

Le **deep learning** est une sous-discipline du ML qui exploite des RN dotés d'architectures multicouches (voir page 15). Les progrès spectaculaires effectués dans ce domaine au cours des dix dernières années ont largement contribué au regain d'intérêt pour les RN et le machine learning en général. Sur le plan conceptuel, on a découvert des astuces qui permettent d'entraîner efficacement ces RN couche par couche (Geoffrey E. Hinton, 2006). Sur le plan du matériel, la disponibilité de processeurs graphiques bon marché (GPU) dotés de plusieurs milliers de cœurs a permis d'accélérer de plusieurs ordres de grandeurs la vitesse d'entraînement. Le DL a permis d'automatiser, d'une manière implicite, la découverte de variables prédictives dans certaines données complexes comme des banques d'images. A ce jour les mécanismes à l'œuvre dans le deep learning, bien qu'efficaces en pratique, restent encore mal compris.

En 2015, combinant astucieusement des techniques de DL et celles d'apprentissage par renforcement, l'équipe de *Google Deep Mind* a conçu un système capable d'apprendre à jouer à une cinquantaine de jeux vidéo *Atari* à un niveau dépassant les meilleurs humains. Le système pouvait observer l'écran pixel par pixel et était informé du score obtenu à l'issue de chaque partie.

Basé sur des techniques similaires, le système *AlphaGo* a battu début 2016 le

champion Lee Sedol au jeu de Go lors d'un tournoi historique. Le point important dans cette percée est qu'*AlphaGo* n'a pas appris à jouer en partant de règles explicites (comme l'avait fait *Deep Blue*) mais en évaluant l'issue de milliers de parties jouées par des humains ainsi qu'en jouant contre lui-même.

L'impression subjective d'un saut quantitatif récent dans les performances de l'IA (pour tout ce qui touche notamment la reconnaissance de formes ou la compréhension du langage parlé) tient moins à des percées conceptuelles majeures qu'au dépassement de certains seuils critique de performance. Ceci a permis de créer des applications qui fonctionnent désormais sans qu'on ait à y penser. Ainsi les taux de reconnaissance vocale qui sont passés en quelques années de 95% à 99% ont rendu possible la réalisation de chatbots qui ne sont plus de simples gadgets.

Pour conclure ce bref survol historique, voici comment Andrew Ng, directeur scientifique chez Baidu, décrit en termes généraux l'état actuel de l'IA (Ng, 2016) :

«Aujourd'hui, en 2016, il est probable que toute tâche mentale qui nécessite moins d'une seconde de réflexion à un humain puisse être automatisée au moyen des techniques d'IA déjà disponibles.»

L'immense majorité des applications de l'IA relèvent aujourd'hui du **machine learning supervisé** (voir page 15) et les progrès récents sont pour l'essentiel à mettre au crédit du Deep Learning.

Un dernier point important à noter est que la principale **source de valeur de l'IA** réside aujourd'hui d'une part dans l'accessibilité, pour un problème spécifique, à de jeux de données de qualité et d'autre part dans la disponibilité de spécialistes qui ont le talent pour optimiser les algorithmes dans un contexte métier particulier. Les algorithmes quant à eux sont pour la plupart accessibles en mode open source.

2.2 LES PERCÉES EN ATTENTE D'UN EINSTEIN DE L'IA

«Imiter la nature ou saisir l'essence de l'intelligence?» Tel est le dilemme fondamental auquel est confronté la recherche en IA depuis ses débuts. Si les RN s'inspirent, même de très loin, de l'architecture du cortex visuel des mammifères, l'approche logiciste se démarque en revanche complètement de la nature. Dans une telle situation il n'est guère surprenant que les problèmes difficiles à résoudre pour une IA soient typiquement d'une nature très différente des problèmes

difficiles à appréhender pour une intelligence humaine.

«Apprendre à cueillir une framboise sans l'écraser est par exemple excessivement difficile pour un robot capable par ailleurs d'évaluer chaque seconde des milliers de configurations de ses membres préhensiles.»

Voici une liste non exhaustive et sans ordre particulier de problèmes qui sont aujourd'hui difficiles ou insolubles pour une IA. Au vu de l'importance que revêt la compréhension du langage naturel, un problème ardu s'il en est, nous y consacrons toute la section 2.3 (voir page 29).

Exploiter de petits jeux de données

La presse IT fait grand cas ces dernières années du big data. De nombreuses organisations disposeraient, sans en être toujours conscientes, de gisements d'informations cachées dans le capharnaüm de données non-structurées accumulées dans leur système d'information. Dans le contexte de l'IA cette intuition est en partie vraie puisque c'est bien la disponibilité d'immenses corpus de textes qui a permis les progrès récents dans la traduction automatique. L'un des principaux enjeux pour l'IA est pourtant de parvenir à utiliser à des fins prédictives de **très petits jeux de données**, ce que les humains font d'ailleurs sans difficulté. Nul besoin en effet de montrer un million d'images de chats ou de tables à un bébé pour qu'il assimile ces concepts.

Peut-être le cerveau d'un enfant y parvient-il au moyen de primitives de modélisation 3D sélectionnées par des millénaires d'évolution pour élaborer avec peu d'informations un modèle physique ou social du monde (*Knight, 2016*). En fait nous n'en savons rien.

Généraliser ou abstraire un apprentissage

AlphaGo est désormais aussi doué qu'un grand maître du jeu de Go. Apprendre un nouveau jeu exigerait cependant un nouvel entraînement du moteur et même, vraisemblablement, une refonte drastique de l'architecture elle-même du système. Un être humain est capable, quant à lui, d'exploiter un apprentissage acquis dans un contexte pour le transférer dans un autre contexte. Aucune technique d'IA ne permet aujourd'hui de reproduire cette faculté **d'abstraction**. Ce point est crucial pour le passage à une IA forte.

Résoudre le problème de l'apprentissage non supervisé

Ce point a déjà été évoqué précédemment, nous le rappelons ici par souci de cohérence. La difficulté à surmonter consiste à faire en sorte qu'une machine soit capable de structurer par elle-même des données complexes. À titre d'exemple : il s'agirait pour un algorithme de parvenir à définir spontanément des concepts de chien, de chat ou d'éléphant sans qu'on ait à lui présenter des centaines de milliers de photos correctement identifiées.



Figure 8 : Les images qui représentent les concepts de visage et de chat appris de manière non supervisée par un réseau profond développé par Google en 2012.

Le DL nous l'avons vu fournit peut-être une esquisse de réponse à cette problématique (Quoc V. Le, 2012).

Apprendre des comportements par observation

L'apprentissage par renforcement, nous l'avons vu, consiste pour un agent intelligent à acquérir un comportement qui le conduira à maximiser une fonction de récompense (ou un bien-être si l'on préfère) au sein d'un environnement complexe. Le problème inverse IRL (**Inverse Reinforcement Learning**), consiste à déterminer (pour un chercheur ou un agent observateur) quelle est la fonction de récompense maximisée par un ou plusieurs agents dont on observe le comportement. Ce problème est beaucoup plus ardu que le premier. Sa résolution est pourtant cruciale si l'on souhaite créer une IA capable d'acquérir un comportement par mimétisme social ou à partir de l'enseignement d'un « maître » qui prodigue des comportements exemplaires. C'est ce que font les humains lorsqu'ils se constituent un système de valeurs par observations de leurs congénères. Résoudre ce problème IRL permettrait alors d'éduquer des agents

⁵ Au rang des spéculations, on peut imaginer qu'une conscience émerge dès lors qu'un ensemble d'agents développent des modèles suffisamment riches à la fois de leur environnement, de leurs interactions avec cet environnement et des relations qu'ils ont avec leurs « semblables ». La conscience pourrait-elle être assimilée à l'identification par un agent qu'une certaine catégorie d'agents de l'environnement sont ses « semblables » ?

intelligents par simple immersion dans la société humaine (dont on supposera le comportement exemplaire pour simplifier...).⁵

Exécuter des tâches mal définies

Imaginons l'ordre suivant adressé à un robot: «Va acheter un cadeau pour l'anniversaire d'Eglantine!» Voilà un exemple de tâche particulièrement mal définie, un problème où l'on parvient d'ailleurs aux frontières des capacités cognitives de la plupart des humains eux-mêmes.

Le cadeau dépend en effet des goûts d'Eglantine qu'on ne connaît pas. Il faut donc se renseigner sur son comportement passé auprès de ceux qui l'ont côtoyé. Une fois cette précieuse information acquise, reste à dénicher un objet qui ait une chance d'augmenter son bonheur... encore un concept mal défini. Enfin, une fois le précieux objet trouvé, encore faut-il s'avoir ou l'acheter et quand. Aujourd'hui, on ne sait pas gérer de tels niveaux d'ambiguïté.

Prendre des décisions avec une rationalité limitée

Les problèmes auxquels un humain est confronté pour survivre dans son environnement sont le plus souvent si complexes qu'une solution parfaitement rationnelle est inenvisageable en pratique. Le temps de délibération nécessaire serait tout simplement prohibitif. Dès lors, l'objectif à atteindre pour une IA ne peut être une rationalité parfaite mais plutôt une rationalité bornée par des ressources de calcul limitées. Bref il s'agit de parvenir à des décisions «assez bonnes» plutôt qu'à des décisions parfaites. Herbert Simon qui a reçu le prix Nobel d'économie en 1978 a développé une théorie descriptive de ce problème de la **rationalité limitée** sans toutefois définir formellement la notion d'être «assez bon».

2.3 LA COMPRÉHENSION DU LANGAGE NATUREL

Pouvoir interagir avec des machines par conversation en langage naturel permettrait d'instaurer entre elles et nous une collaboration d'une efficacité sans commune mesure avec celles des **IHM** actuelles. Il suffit pour cela d'imaginer un système capable non seulement de prendre une décision pertinente mais encore d'expliquer pourquoi il l'a prise. Nul doute qu'il s'agirait là d'une révolution technologique majeure. Aujourd'hui, hélas, nous en sommes encore très loin. Ainsi personne ne sait concevoir un système qui serait capable d'écouter des

conversations durant deux ans tout en apprenant progressivement à parler.

Certains chercheurs estiment d'ailleurs que le problème de la compréhension du langage naturel est **AI-complet**, signifiant par-là que sa résolution impliquerait, de fait, la résolution du problème central de la création d'une AI forte (*Knight, 2016*).

Comparés aux balbutiements d'*ELIZA* il y a un demi-siècle les progrès de chatbots comme *Siri*, *Cortana* ou *Google Now* sont certes significatifs. Ils reposent pourtant moins sur une authentique compréhension de langage naturel, au sens où nous le souhaiterions, qu'aux progrès récents dans la reconnaissance de la parole qui sont à mettre au crédit du deep learning. En schématisant un peu, ces chatbots ne font rien d'autre que mettre en correspondance une liste d'actions prédéfinies (pour nos smartphones) avec une liste de questions identifiées par avance. A proprement parler ces chatbots ne comprennent strictement rien !

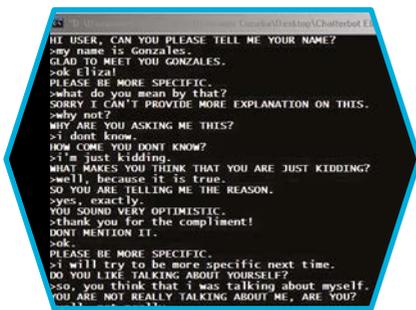


Figure 9 : Un demi-siècle de progrès modestes dans la compréhension du langage naturel : *ELIZA* en 1964, *Siri* en 2015.

Comme nous l'avons déjà mentionné, l'approche statistique à la traduction automatique a apporté une solution partielle au problème de désambiguïsation de termes selon le contexte. Pour autant, elle ne constitue qu'un pas très modeste vers la compréhension du sens exprimé dans un texte.

Un exemple d'algorithme, mystérieux par son efficacité dans l'analyse du langage, est l'algorithme *Word2Vec* créé par *Google*. Après avoir ingurgité un corpus de texte *Word2Vec* crée un espace de plusieurs centaines de dimensions dans lequel chaque mot est représenté par un point, et ceci de telle sorte que des mots qui partagent des contextes sémantiques proches soient également voisins dans cette représentation. Il s'avère que cette représentation vectorielle permet d'identifier des analogies sémantiques entre différents mots. A partir d'un couple de mots comme (« Paris », « France ») et du mot « Italie » on pourra déduire

«Rome» par une simple opération arithmétique sur les vecteurs représentatifs :
 Paris - France + Italie = Rome !

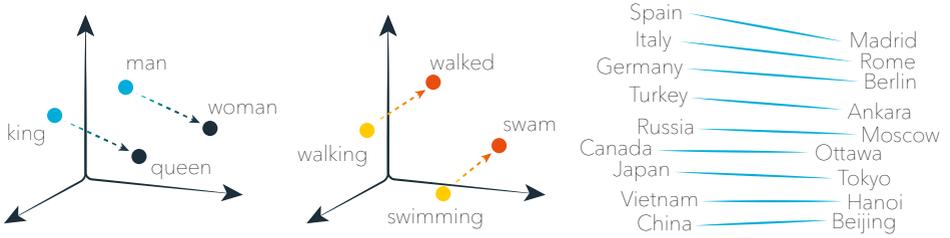


Figure 10 : L'algorithme Word2Vec permet d'associer des points dans un espace à grandes dimensions à des mots de manière à permettre d'identifier des analogies.

Existe-t-il des impossibilités de principe à une solution complète du problème de la compréhension du langage naturel? Comprendre une langue exige-t-il une conscience? A ce jour ces questions restent ouvertes.



3

LES USAGES

3.1 AUJOURD'HUI

Enumérer une liste exhaustive des applications actuelles de l'IA s'apparente rapidement à faire un inventaire à la Prévert, tant elles sont déjà nombreuses et variées. Aussi, nous nous limiterons à quelques exemples parmi les plus significatifs. Beaucoup de ces applications ont bénéficié des progrès récents dans le domaine du Deep Learning (voir page 24) comme nous l'avons déjà indiqué.

Les systèmes de recommandation

Dès le début des années 2000 *Amazon* a popularisé les systèmes de recommandation avec sa célèbre rubrique «les clients qui ont acheté cet article ont aussi achetés...». Les techniques de **filtrage collaboratif** sont aujourd'hui les plus utilisées. Elles exploitent des modèles statistiques capables d'estimer des affinités entre des utilisateurs et des produits à partir d'un grand nombre d'exemples. Chaque utilisateur se voit proposer une liste de produits susceptibles de de l'intéresser. Cette fonctionnalité est désormais devenue vitale pour toutes les plateformes de vente en ligne qui proposent d'énormes catalogues de produits comme *Netflix* ou *Spotify* et également pour les réseaux sociaux comme *LinkedIn*.

Ces systèmes de recommandation découvrent par eux-mêmes des appariements, parfois insoupçonnés, entre des utilisateurs et des produits. A ce titre ils peuvent être envisagés, si l'on veut, comme une forme **rudimentaire d'IA** au même titre par ailleurs que tous les systèmes qui exploitent les techniques de machine learning.

Les assistants pour smartphone

Les chatbots tels que *Siri*, *Cortana* ou *Google Now* combinent des fonctionnalités de reconnaissance vocale avec une esquisse de traitement du langage naturel pour rendre plus naturelle nos interactions avec nos smartphones. Ils exploitent également des informations contextuelles, géographiques et sociales, pour affiner leurs actions. Ces assistants ne se contentent pas de répondre à des questions mais font également des suggestions contextuelles. Ils s'améliorent et se personnalisent au fur et à mesure que l'utilisateur interagit avec eux.

Les progrès récents en matière de reconnaissance vocale, passée en quelques années d'une fiabilité de 95% à 99% grâce notamment au deep learning (DL), ont largement contribué à populariser ces chatbots.

Garantir la **sécurité** de ces systèmes est aujourd'hui l'un des principaux challenge à relever pour ces technologies. L'entraînement des modèles de DL nécessite en effet de grandes quantités de données qui ne peuvent être traitées localement par le CPU d'un smartphone. Deux approches complémentaires sont actuellement utilisées pour garantir la sécurité :

- Les techniques cryptographiques de **Differential Privacy** (exploitée notamment par *Apple* à partir d'*iOs 10*) permettent de traiter à distance de grandes quantités de données tout en garantissant mathématiquement qu'aucune donnée individuelle ne pourra être extraite à partir des données transmises.
- Une autre technique qui fait encore l'objet de nombreuses recherches est l'**encryption homomorphe**. Elle permet de déléguer des traitements de manière totalement anonyme en opérant directement sur les données cryptées elle-même, sans qu'il soit nécessaire à aucun moment de les décrypter.

La reconnaissance d'images

Les progrès du DL permettent aujourd'hui une reconnaissance très efficace des visages. Ces techniques sont couramment exploitées dans les systèmes de vidéosurveillance, les réseaux sociaux et les logiciels domestiques de classification de photos. Des outils existent qui permettent de classer des images par type de scènes avec un niveau de précision suffisant pour autoriser l'exploration de grandes bases d'images.



Figure 11 : Classification automatique des scènes par Google.

En 2015 sont apparus les premiers **systèmes** capables de décrire en texte clair le contenu d'une image, (*Andrej Karpathy, 2015*). Bien que les descriptions ne soient pas dépourvues d'erreurs, les performances de ces algorithmes sont tout

de même impressionnantes.

On peut envisager ces derniers systèmes comme des tentatives de compréhension du monde par des IA au même titre que celles qui visent à comprendre le langage naturel. A partir de quelques dizaines de milliers d'images décrites chacune par plusieurs descriptions en texte clair, il est possible d'entraîner des RN conçus à cet effet pour qu'ils génèrent des descriptions crédibles de nouvelles images. Les techniques utilisées exploitent astucieusement différents types de RN pour corréliser la structure syntaxique des descriptions textuelles avec le contenu des images. Superficiellement tout se passe donc comme si le système « comprenait » ce que représente l'image. Insistons sur le fait qu'il ne s'agit pas de simples objets ou des personnes qui sont reconnus mais bien des situations et des actions ce qui en fait probablement l'une des applications les plus spectaculaire de l'IA à l'heure actuelle.

S'agit-il pour autant d'un premier pas vers l'élaboration d'un modèle du monde par des machines? Probablement pas. Un enfant n'a pas besoin de revoir des milliers de fois une situation pour l'identifier correctement. Le cerveau humain s'appuie vraisemblablement sur des primitives qui lui permettent de modéliser très efficacement des modèles physiques de monde (Knight, 2016).

DESCRIBES WITHOUT ERRORS



A person riding a motorcycle on a dirt road



A group of young people playing a game of frisbee



A herd of elephants walking across a dry grass field

DESCRIBES WITH MINOR ERRORS



Two dogs play in the grass



Two hockey players are fighting over the puck



A close up of a cat laying on a couch

SOMEWHAT RELATED TO THE IMAGE



a skateboarder does a trick on a ramp



A little girl in a pink hat is blowing bubbles



a red motorcycle parked on the side of the road

UNRELATED TO THE IMAGE



A dog is jumping to catch a frisbee



A refrigerator filled with lots of food and drinks



A yellow school bus parked in a parking lot

Figure 12: Exemples de description automatique d'images par Google au moyen de réseaux de neurones récurrents.

La traduction automatique

Bien que les systèmes de traduction automatique n'en soient pas encore à traduire La Pléiade en mandarin, ils peuvent cependant rendre intelligible une page d'un journal écrit dans une autre langue. *Google* propose par ailleurs aux internautes de se fédérer en communauté pour améliorer progressivement les traductions proposées ce qui est une caractéristique fondamentale des systèmes apprenants.

Les **méthodes statistiques** utilisées aujourd'hui exploitent de gigantesques corpus de documents mais éludent complètement le problème profond de l'accès à la sémantique des textes. Ces systèmes ne comprennent pas à proprement parler ce qu'ils traduisent, ils se contentent d'exploiter des corrélations entre des groupes de mots dans différentes langues.

La **sémantique** d'un texte s'apparente plutôt une sorte de corrélation, ou de mise en correspondance, entre des structures syntaxiques d'une langue et des situations du monde réel. Y accéder exigerait par conséquent de construire une représentation du monde dans son ensemble qui ne peut vraisemblablement être obtenue que par interaction avec le monde physique. Il est dès lors concevable que les systèmes de traduction tels qu'ils sont conçus aujourd'hui se heurtent un jour à une limite infranchissable dans la qualité des traductions qu'ils peuvent proposer. Pour autant cela ne les empêchera pas d'être utiles en pratique.

Applications dans l'industrie du divertissement

Dans les jeux vidéo on utilise l'IA pour conférer une personnalité à des personnages dont le comportement doit tenir compte de leurs interactions avec l'environnement et avec d'autres agents présents dans le jeu. On utilise également des RN pour la génération de textures d'aspects aléatoires dans des décors synthétiques. Enfin l'IA est utilisée également pour analyser le comportement des joueurs et améliorer la jouabilité.

Application dans les transports

Nos GPS sont des systèmes de planification largement répandus. Les systèmes embarqués dans les modèles de véhicules haut de gamme sont d'ores et déjà capables de reconnaître certaines situations à risque dans le trafic et de prendre les mesures préventives qui s'imposent (freinage d'urgence). Le concept de **voiture sans conducteur** (voir page 38), est encore en cours de développement mais des résultats encourageants sont annoncés pour la *Google Car*, 17 accidents

mineurs sur 3.2 millions de kilomètres parcourus, ainsi que pour les véhicules développés par *Tesla*.

Robotique domestique

La robotique domestique se heurte pour l'instant à la difficulté de construire des systèmes mécaniques à la fois flexibles et robustes. Pour l'heure, la seule application convaincante de l'IA dans ce secteur reste celle des **aspirateurs intelligents**. Capables de construire un modèle 3D de leur environnement grâce à un système de vision artificielle ils optimisent de manière autonome leurs déplacements dans un appartement.

3.2 BIENTÔT

L'exercice d'anticipation est toujours risqué, quel que soit la technologie, mais dans le cas de l'IA il est plus délicat encore car il est bien difficile en effet d'imaginer un secteur d'activité qui ne bénéficierait pas, directement ou indirectement, d'un surcroît d'intelligence. Il y a fort à parier que d'ici quelques années fleuriront des myriades de startups qui ajouteront un élément d'IA à n'importe quelle activité existante.

Des tâches d'expertise administrative, juridique, médicale ou technique pourront être prises en charge par des IA spécialisées fonctionnant en mode collaboratif avec les humains qu'elles assisteront et avec qui elles interagiront sur un mode conversationnel. Les humains se consacreront aux tâches qui exigent un jugement (éthique ou transdisciplinaire), de l'imagination, des facultés d'abstraction ou encore celles qui demandent une capacité d'empathie seule apte à créer du lien social.

L'IA sera-t-elle l'électricité de demain? Ce n'est pas impossible si l'on songe à son potentiel disruptif dans presque toutes les activités humaines qu'il s'agisse d'urbanisme, de santé publique, de transports, d'emploi, de formation ou de divertissement.

Nous nous limiterons cependant ici à quelques extrapolations prudentes pour les années à venir à partir des usages existants.

Les voitures sans conducteur

La voiture individuelle sans conducteur, dont la *Google Car* est aujourd'hui le prototype le plus connu, sera l'une des applications les plus visibles de l'IA ces quinze prochaines années. La diminution attendue des accidents de la route (on évoque un facteur 100) contribuera à changer profondément la perception du grand public face à l'IA puisque celle-ci aura en charge aussi bien la sécurité des passagers d'un véhicule que l'optimisation globale du trafic urbain.

Les impacts sociaux-économiques de ce nouveau mode de **transport individuel automatisé** peuvent difficilement être surestimés. L'urbanisme des grandes agglomérations en sera chamboulé avec la possibilité de redistribuer le temps de travail. Les voitures conduites par des humains seront rapidement perçues comme dangereuses et probablement interdites après une période transitoire. Le secteur des assurances sera lui aussi affecté puisque la responsabilité en cas d'accidents n'incombera plus au possesseur du véhicule mais à son concepteur.



Figure 13 : Le prototype de voiture sans conducteur par Google.

La relation affective à la voiture comme attribut d'un statut social qui prévaut depuis bientôt un siècle disparaîtra progressivement. Acheter une voiture pourrait devenir superflu pour les particuliers, du moins dans les grandes agglomérations. Un véhicule automatique aurait alors la même charge émotionnelle qu'un ascenseur aujourd'hui.

Une nouvelle médecine

Alors que la voiture sans conducteur nous promet d'éradiquer les accidents de la route, la médecine assistée par l'IA rend crédible à moyen terme l'accès

⁶ Artificial Intelligence in Medicine: aimjournal.com

quasi universel à une médecine préventive de pointe et ultra-personnalisée. Des revues spécialisées sont consacrées à ce sujet⁶.

La relation patient-médecin est immuable depuis des millénaires: un patient décrit une souffrance à laquelle le médecin répond par un diagnostic puis une proposition de traitement. Il prend ses décisions en s'appuyant à la fois sur ses connaissances académiques (ou ancestrales), sur l'expérience glanée auprès de patients aux symptômes similaires et sur ce qu'il sait des antécédents de son patient. L'AI et le machine learning peuvent intervenir à toutes les étapes de cette relation :

- La compréhension de la description des symptômes par le patient
- La prise en compte de l'intégralité du savoir médical consigné dans les livres et les revues médicales
- L'analyse comparative des symptômes et de l'issue des traitements pour tous les cas similaires connus
- L'enrichissement en continu d'une base de symptômes et de traitements
- La personnalisation et le suivi rigoureux et en temps réel d'un traitement grâce à de nouveaux objets connectés

«Une IA-chatbot ne saurait évidemment se substituer à la relation interpersonnelle entre un patient et son médecin. Il faut plutôt l'envisager comme un assistant au diagnostic ou comme un interlocuteur capable d'orienter un patient vers le spécialiste adéquat.»

Les startups (Furness, 2016) qui se lancent dès aujourd'hui dans le diagnostic assisté par IA motivent leur démarche par la **réduction drastique des déficits des systèmes de santé** qu'impliquerait, selon eux, ce nouveau mode d'accès à l'expertise médicale. De nombreuses consultations inutiles pour de petits bobos pourraient ainsi être évitées. La prévention des maladies mentales est un domaine où l'IA apporte des atouts considérables car les expériences déjà menées démontrent que, sur ces questions intimes, les patients se livrent plus aisément à une machine qu'à un médecin.

Mentionnant encore l'utilisation des techniques de deep learning qui, par les corrélations qu'elles mettent en évidence entre structures moléculaires et effets pharmacologiques, permettent l'élaboration de **nouveaux médicaments**.

Un nouveau type d'IHM : la conversation

À mesure que se perfectionneront les assistants vocaux comme *Siri*, *Google Now* ou *Cortana*, se développera progressivement une nouvelle manière d'interagir avec les machines, plus naturelle : la conversation. Aujourd'hui limités à quelques interactions stéréotypées, ces assistants seront bientôt capables de faire connaissance avec leur interlocuteur et de prendre en considération le contexte émotionnel, social et géographique dans lequel il s'exprime. L'IA offrira à tout un chacun, et notamment aux personnes de ressources modestes, un **service de conciergerie** aujourd'hui réservé à quelques businessmen surbookés.

Les experts (*Ng, 2016*) prédisent que d'ici quelques années 50% de nos interactions avec les machines se feront sur un mode conversationnel.

Une technologie qui sait (enfin) se faire oublier

Hypnotisés jusqu'à l'incivilité par les sollicitations incessantes de nos smartphones, nous nous laissons trop souvent envahir par des technologies pourtant conçues à l'origine pour simplifier notre quotidien. Courrier, agendas, messageries, réseaux sociaux nous bombardent à chaque instant de leurs notifications, ceci au détriment de notre capacité d'attention aux autres et à nous-même.

Une vision optimiste de l'IA consiste à envisager que celle-ci pourrait contribuer à réhabiliter une valeur trop longtemps oubliée de notre civilisation agitée : **la tranquillité** ! Des technologies d'IA équipant nos futurs smartphones pourrait ainsi contribuer à fluidifier nos vies en les débarrassant d'une part de leur agitation superflue, soit par anticipation de nos besoins, soit par filtrage des sollicitations au moyen de données contextuelles. C'est par exemple la vision portée par une startup comme *SNIPS* dont l'objectif avoué est de faire « disparaître la technologie » !

Dans cette perspective, très optimiste reconnaissons-le, l'IA pourrait alors contribuer à réorienter l'idée même de progrès civilisationnel défini désormais comme une reconquête. Celle d'un temps et d'une liberté que l'on pourrait désormais consacrer à l'essentiel d'une vie d'homme : aller à la rencontre des autres, s'émerveiller des beautés du monde, découvrir les grands penseurs, créer un œuvre pour ceux qui en ont le talent.

Les moteurs de recherche au service de l'IA

Les technologies d'IA ont permis à *Google* d'améliorer son moteur de recherche

en injectant une esquisse d'interprétation sémantique aux documents indexés. Il est possible que le futur proche voie cette relation entre IA et moteur de recherche s'inverser. D'ici quelques années le principal produit de *Google* pourrait bien ne plus être un moteur de recherche mais plutôt une IA alimentée par les milliards de requêtes et de clics quotidien des internautes.

3.3 DEMAIN ?

Le futur de l'IA est un lieu peuplé de fantômes. Dans notre époque anxieuse, nombre de ces projections ne dévoilent en réalité que des craintes face à nos propres limitations : qu'advierait-il en effet si des machines omniscientes et omnipotentes se trouvaient animées d'aussi mauvaises intentions que celles que nous nourrissons à l'égard de nos semblables ?

Bien qu'il n'existe à ce jour aucune roadmap crédible pour la création de machines hyper-intelligentes dotées de conscience et de volonté, voici quelques-unes des questions que soulèverait leur avènement :

- Des systèmes autonomes dotées de conscience ne seraient plus, en vérité, des machines. Se poserait dès lors la question des droits qu'il conviendrait de leur accorder. Après les droits de l'homme, il faudrait d'élaborer un **droit des machines intelligentes**. Les deux législations devraient-elles être différentes ? Une relation de subordination aux humains serait-elle légitime d'un point de vue éthique ?
- Peut-on concevoir des **machines hyper-intelligentes** tout en garantissant que celle-ci n'agiront que dans l'intérêt de l'humanité ? Existe-t-il une contradiction fondamentale ou un conflit éthique entre l'ambition de créer une IA autonome et l'assujettissement dans lequel nous souhaiterions la maintenir pour notre propre bénéfice ?
- La première machine hyper-intelligente jamais créée serait capable de s'améliorer mieux que nous ne saurions le faire. À partir de ce point, que certains nomme la **singularité technologique**, le progrès ne serait plus le fait des hommes mais celui des machines. Faut-il nous en inquiéter ou y voir au contraire un nouveau chapitre de l'évolution dont l'humanité serait la chrysalide ?

Nous laissons le reste à l'imagination fertile des auteurs de science-fiction.



UNE ÉTHIQUE ET UN DROIT POUR L'IA

Face au potentiel disruptif incommensurable de l'IA, l'objectif d'une éthique adaptée à cette nouvelle donne technologique est simple à formuler: elle doit nous permettre discerner où se trouve l'intérêt général en évitant aussi bien l'écueil des **peurs infondées** et que celui d'un **techno-scientisme béat** qui voudrait nous convaincre que le progrès technologique à lui seul viendra à bout de tous les maux qui affligent l'humanité.

Certaines questions juridiques soulevées par l'IA sont inédites comme nous le verrons. Les questions éthiques en revanche sont peu ou prou les mêmes que celles qui se posent à toute société qui cherche à concilier démocratie et innovation technologique :

- Où se trouve l'**intérêt général** lorsqu'une innovation ouvre un nouveau champ de possibles ?
- Comment arbitrer entre **la nécessité de maîtriser les risques** que fait naître une technologie et les **coûts** qu'engendre cette maîtrise ?
- Comment arbitrer entre les **intérêts privés**, notamment la préservation de la propriété intellectuelle des investisseurs, et les exigences d'un **contrôle démocratique** ?
- En tant que technologie de l'information, l'IA est-elle susceptible d'affecter certains attributs essentiels de notre humanité comme notre capacité de jugement ou notre libre-arbitre ?

«L'éthique, comprise ici comme un ensemble de valeurs, ne formule naturellement aucune réponse directe à ces questions. Elle n'est que l'aiguillon qui doit initier le débat politique et citoyen autour de la question fondamentale qui est de savoir quel futur nous souhaitons construire à l'aide de ces nouvelles technologies.»

Parmi les toutes questions éthiques que soulève l'IA en voici cinq qui nous paraissent importantes.

4.1 COMMENT RÉPARTIR ÉQUITABLEMENT LES RICHESSES CRÉÉES PAR L'IA ?

Peu importe la nature d'une activité économique, celle-ci pourra tirer bénéfice des technologies d'IA à partir du moment où la création de valeur fera intervenir un processus cognitif automatisable. Les coûts de production de nombreux

biens et services vont par conséquent diminuer. Mais, **à qui alors profitera cette nouvelle manne**? Aux concepteurs des systèmes d'IA qui auront permis ces économies? Aux acquéreurs de ces systèmes? A la société dans son ensemble? Si ces technologies et les données qu'elles utilisent devaient rester confinées dans les mains de quelques privilégiés, les inégalités de richesses déjà importantes dans nos sociétés pourraient s'en trouver amplifiées pour parvenir à un point de rupture où elles pourraient menacer leur stabilité même.

Comme toute nouvelle technologie l'IA **détruira des emplois et créera simultanément de nouvelles opportunités**. Dans un premier temps il est plausible que l'IA ne supprimera que des fonctions plutôt que des emplois. Les tâches qui disparaîtront les premières seront toutes celles qui ne relèvent que d'un traitement celles qui ne relèvent que d'un traitement de l'information mais n'exigent ni jugement, ni créativité, ni empathie envers d'autres personnes. Dans les métiers commerciaux par exemple les humains pourront dès lors se concentrer sur la relation client plutôt que sur l'expertise technique désormais déléguée aux IA. Il en ira vraisemblablement de même dans les métiers de la santé, du droit et de l'administration: les IA fourniront une expertise technique ultra-pointue, systématiquement mise à jour, à des professionnels dont la valeur ajoutée ne se situera pas dans l'expertise mais dans la qualité des relations entretenues avec leurs clients, leurs partenaires ou leurs patients.

«Dans un monde où l'expertise technique serait quasiment gratuite et accessible à tous, la différence entre deux prestations se ferait essentiellement sur la confiance et la civilité des relations entre partenaires.»

En somme, voici une perspective plutôt réjouissante!

De nouveaux métiers très qualifiés apparaîtront vraisemblablement pour quelques experts en IA dans le sillage des métiers actuels de la data science. Le bilan en termes d'emplois est toutefois difficile à anticiper car il est plus aisé de prévoir les disparitions d'activités que les nouvelles opportunités. À moyen terme, les pouvoirs publics auront la responsabilité de proposer une reconversion ou une indemnisation aux personnes dont les emplois auront disparu avec la généralisation de ces technologies. A plus long terme les gains de productivité apportés par l'IA pourraient même nous contraindre à repenser de fond en comble notre rapport au travail et le modèle de fiscalité qui lui est associé.

La recherche de pointe en IA est aujourd'hui largement effectuée par les grandes entreprises américaines du web (et par l'armée américaine⁷), les **GAF**A pour ne pas les nommer. Certaines d'entre elles voient leur intérêt dans une large ouverture des technologies qu'elles développent. C'est le cas de *Google* par

⁷ Le revenu inconditionnel de base est le plus connu de ces nouveaux modèles.

exemple, qui a ouvert son framework de machine learning *TensorFlow* ou de *Facebook* qui propose son framework *Torch* pour le Deep Learning. D'autres au contraire, comme *Apple* par exemple, sont moins enclines au partage.

Garantir que l'intérêt général reste une priorité exigerait cependant qu'une part significative de cette **recherche en IA** soient menée **sur des fonds publics** et sans contraintes de rentabilité immédiate. Il est donc essentiel que les pouvoirs publics, à tous les échelons de décision (*Dadich, 2016*) soient sensibilisés aux enjeux de l'IA, non seulement sur ces questions économiques mais aussi sur les aspects juridiques que nous évoquons dans la section 4.2 et 4.3.

4.2 COMMENT GARANTIR LA NEUTRALITÉ ET LA LÉGALITÉ D'UN SYSTÈME PRÉDICTIF ?

De plus en plus d'algorithmes de machine learning (qui relèvent par conséquent de l'IA au sens large) formulent des recommandations susceptibles d'avoir un impact majeur sur nos vies. Qu'il s'agisse d'évaluer des compétences dans le cadre d'un processus de recrutement, de prédire les chances de récidive d'un délinquant avant une décision d'incarcération ou encore de mesurer un risque de crédit avant l'octroi d'un prêt. Dans tous ces cas se pose la question de la **neutralité des algorithmes et des données** qui les alimentent. En d'autres termes il est important de pouvoir garantir qu'aucun facteur de discrimination incompatible avec un état de droit ne vienne se glisser dans l'élaboration d'une prédiction ou d'une recommandation.

Insistons cependant sur un point : ces dangers, bien réels, ne doivent pas occulter le fait que l'IA utilisée à bon escient pourrait au contraire contribuer dans bien des situations à améliorer l'**impartialité** de décisions entachées jusque-là par des préjugés très humains !

Les facteurs d'iniquités sociales peuvent avoir différentes sources que nous examinons brièvement dans les trois paragraphes qui suivent.

Des algorithmes avec des biais statistiques

Rappelons qu'un **bias statistique** désigne une caractéristique d'un algorithme qui fournit une estimation (ou une prédiction) systématiquement erronée de certaines caractéristiques d'une population ou d'un historique d'événements.

L'une des difficultés pour parer à ce genre de défaut tient au fait que les

algorithmes sont des objets mathématiques, à la fois opaques et enduits d'un vernis d'objectivité scientifique (O'Neil, 2016). Ceci peut leur conférer un caractère intimidant propre freiner leur examen critique par des non spécialistes.

La détection et la correction de biais statistiques d'algorithmes prédictifs devra par conséquent être confiée à des chercheurs et des journalistes spécialisés qui en inspecteront scrupuleusement le fonctionnement. Cette **exigence d'ouverture** des systèmes d'IA devrait être garantie par la loi, même s'ils sont conçus par des sociétés privées.

Des données qui encapsulent des opinions

Des **facteurs discriminants** peuvent cependant être introduits dans une prédiction même lorsque l'algorithme utilisé n'est entaché d'aucun biais statistique. C'est le cas lorsque les données qu'il exploite reflètent implicitement certains présupposés idéologiques. Ainsi un système d'aide aux décisions d'incarcération qui s'appuierait sur un historique d'arrestations qui reflèterait, même partiellement, certains comportements xénophobes d'une partie des forces de l'ordre n'a évidemment aucune chance d'être équitable (Loukides, 2016).

«La notion de biais statistique, qui caractérise un algorithme, ne doit en aucun cas être confondue, comme c'est trop souvent le cas, avec la notion plus générale de facteur de discrimination qui peut être induit par des données entachées de présupposés idéologiques.»

Des données prédictives sans valeur légale

Imaginons un système prédictif qui, pour évaluer la dangerosité d'un individu, tiendrait compte de sa commune de résidence ou des antécédents judiciaires de certains membres de sa famille. Un tel système pourrait, là encore, introduire des facteurs de discrimination inacceptables dans un état de droit, et ceci, quand bien même aucun biais algorithmique ni aucune donnée idéologiquement connotée ne seraient à mettre en cause.

En termes généraux la question peut se formuler de la manière suivante : est-il légitime d'utiliser toutes les données qui ont un caractère prédictif dès lors que celles-ci sont neutres idéologiquement ? Même s'il existe probablement un **vide juridique** en la matière aujourd'hui, un principe de bon sens voudrait qu'il soit interdit d'exploiter à des fins prédictives toute donnée qui n'aurait par ailleurs aucune valeur légale devant un tribunal.

La reconnaissance d'images à grande échelle pourrait être exploitée par exemple pour croiser des informations récupérées auprès de sources diverses comme les réseaux sociaux pour désanonymiser certaines informations. Là encore il existe probablement un vide juridique qu'il faudra combler en connaissance de cause.

4.3 QUELLES RESPONSABILITÉS EN CAS D'ACCIDENT ?

A partir du moment où des IA interviendront dans l'**initiation de certaines transactions financières**, dans le **choix de la trajectoire d'un véhicule** face à un danger imminent ou dans la décision d'**engager une arme** des accidents surviendront, inévitablement. Qui alors sera responsable ? L'utilisateur du système ? Son concepteur ? L'autorité de certification qui aurait échoué à anticiper certains dysfonctionnements ?

Des difficultés juridiques inédites surviendront. Comment en effet donner un sens à la notion de « préjudice prévisible » à laquelle font référence nombre de textes légaux alors même que les comportements d'une IA pourraient être, par nature, imprévisibles ? Qui aura la charge d'anticiper ces dysfonctionnements d'une IA ? Autant de questions qui n'ont pas de réponses simples. Probablement est-il impossible d'anticiper l'impact de l'IA sur la législation actuelle dans tous les domaines mais le législateur devrait y réfléchir dès aujourd'hui, au cas par cas.

Dans le cas des véhicules automatiques, certains choix éthiques difficiles devront être fait en anticipation de situations potentiellement létales auxquelles une IA aura à réagir. En cas d'accident imminent quelles vies sauver ? Celle du passager-acquéreur du véhicule automatique ? Celle d'un groupe de personnes mises en danger ?

Il y a fort à parier que tout accident impliquant une IA défraiera la chronique, quel que soient par ailleurs le niveau de fiabilité des systèmes. Assumer soi-même un risque relativement élevé est en effet psychologiquement nettement plus acceptable que de déléguer la gestion d'un risque, même minime, à un système automatique. L'intérêt général de sécurité pourrait en l'occurrence entrer en conflit avec nos instincts primitifs de survie individuels.

Face à ces risques inédits se profile la **tentation d'une réglementation excessive** qui pourrait se faire au détriment des capacités d'innovations des entreprises et du développement de leur sens des responsabilités. Un point sur lequel une vigilance toute particulière est de mise, en particulier chez nous en France.

4.4 UNE NOUVELLE FORME D'ASSISTANAT ?

Le débat est récurrent à chaque innovation majeure dans les technologies de l'information: quel sera son impact sur nos propres capacités cognitives? L'un des premiers exemples remonte sans doute à Platon avec son [discours sur les dangers de l'écriture](#)⁸ qui introduirait selon lui «l'oubli dans l'âme de ceux qui l'utilisent». Dans les années 1970 l'avènement des premières calculatrices dans les écoles déclencha des débats houleux entre pédagogues et suscita des inquiétudes chez les parents: nos enfants n'apprendront-ils désormais plus à calculer? Plus près de nous, le célèbre article de Nicolas Carr, *«Is Google Making us Stoopid?»* fit également grand bruit. Son argumentation reposait sur l'idée que l'acquisition d'information par le web aurait à long terme un effet abrasif sur notre capacité d'attention et notre aptitude à la réflexion concentrée. Bien que cette argumentation ne fût à l'origine étayée par aucune étude psychologique ou neurologique à long terme, et pour cause, elle eut un large écho car elle coïncidait avec l'intuition et l'expérience de beaucoup d'internautes.



Figure 14: A chaque nouvelle techno IT, une même question: quel impact sur nos cerveaux? L'écriture, la TI-30, le web, l'IA.

Gageons qu'une nouvelle instance de ce débat resurgira lorsque les outils d'aide à la décision basés sur des technologies d'IA se seront plus largement démocratisés. Si l'interaction avec le web «à l'ancienne» suscite des craintes légitimes face à la surcharge informationnelle qu'il inflige à nos cerveaux, l'IA généralisée pourrait à l'inverse poser la question de l'impact d'une **assistance excessive**. L'IA nous promet de manière crédible de nous assister dans un nombre croissant de processus de décision, d'évaluation, d'organisation et de tri de l'information. Dans toutes ces situations notre intuition pourra être remplacée avec profit par des analyses factuelles reposant sur le machine learning. Ressurgit alors la vieille appréhension de Platon, celle d'un ramollissement insidieux de notre intellect qui, tel un muscle manquant d'exercice risquerait de s'atrophier progressivement.

⁸ On relira avec intérêt l'argumentation de Platon dans *Phèdre* qui possède une modernité troublante lorsqu'on fait le parallèle avec l'article de Nicolas Carr, *Is Google Making us Stoopid?*

Comme toujours, la vigilance éclairée sera meilleure conseillère que la simple crainte.

4.5 ET SI L'HOMME SE RAPPROCHAÎT DES MACHINES (PLUTÔT QUE L'INVERSE)?

Toute l'ambition de la recherche en IA repose en définitive sur le postulat que les performances cognitives des machines pourront se rapprocher progressivement de celles des humains. A bien y réfléchir, ce rapprochement pourrait cependant intervenir, à long terme, de deux manières. La première hypothèse, optimiste, présuppose que les machines verront leurs performances augmenter. Mais, au risque de jouer les Cassandre, on pourrait imaginer aussi que celles des humains déclinent par suite d'un mimétisme inconscient consécutif à une interaction prolongée avec des IA aux performances conversationnelles limitées.

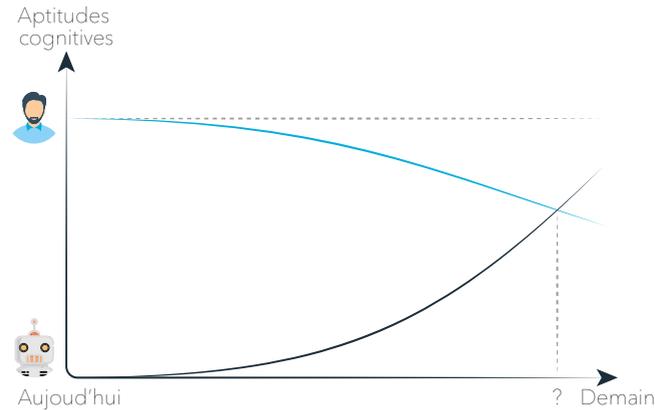


Figure 15: Quelle évolution pour les aptitudes cognitives humaines?

Imaginons que nous interagissions de plus en plus par conversation avec des machines aux facultés de compréhension limitées, des machines naturellement dépourvues de tout sens de l'ironie, d'humour, d'abstraction, d'ellipse ou de périphrase. Pour nous faire comprendre d'elles nous pourrions alors inconsciemment consentir à appauvrir notre propre expression orale. A l'image du très pauvre et très inélégant franglais managérial pratiqué dans nos réunions et nos conférences.



5

**L'IA ET LE MÉTIER
DU CONSEIL**

Dans cette dernière partie nous souhaitons brièvement aborder la question de l'impact des technologies d'IA, présentes ou avenir, sur le **métier du conseil** tel que nous le pratiquons chez *Weave* : à savoir l'accompagnement des entreprises dans leurs choix technologiques et dans la transformation de l'organisation qu'ils impliquent.

Afin de dissiper toute ambiguïté, précisons d'emblée que l'on pourrait envisager deux questions distinctes. La première consisterait à anticiper les usages possibles de l'IA pour les consultants en charge d'une mission de transformation classique. Face à un problème d'organisation particulièrement complexe ou en présence de contraintes réglementaires inextricables, on pourrait concevoir que des consultants aient recours pour accomplir leur mission à l'expertise d'une ou plusieurs IA. Ainsi pourrait-on imaginer par exemple que toutes ou partie des tâches administratives et de planification d'un PMO soient un jour prises en charge par une IA (Vegard Kolbjørnsrud, 2016). Pour intéressante qu'elle soit, cette question nous semble pour l'heure toutefois trop spéculative pour être abordée sérieusement. Aussi, nous nous intéresserons plutôt à la problématique plus immédiate qui consiste à accompagner une entreprise ou une direction des systèmes d'information (DSI) dans l'automatisation de certains processus cognitifs au moyen des technologies existantes d'IA.

A l'échelle d'une l'entreprise l'étude d'impact des technologies d'IA interviendra en amont des projets technologiques, typiquement durant l'élaboration d'un **schéma directeur**. Quels processus métiers, quelles expertises pourraient bénéficier, à court ou moyen terme, de de l'automatisation de processus cognitifs comme l'apprentissage d'un corpus de documents ou l'interprétation de requêtes formulées en langue naturelle ? Quels seront les impacts sur le recrutement, sur les besoins de formation ou de reconversion des collaborateurs consécutifs à la mise en œuvre d'une IA ? Existe-t-il des risques liés à un éventuel vide juridique ? Telles sont quelques-unes des questions qu'il conviendra d'aborder.

Ces évolutions laissent présager l'importance pour une société de conseil de faire la preuve de sa capacité à travailler au sein d'un **écosystème de startups** dont elle aura à coordonner les interventions, bien souvent dans des **contextes culturels peu enclins à l'agilité et à l'expérimentation**. Elle devra démontrer par ailleurs son aptitude à mener une **veille technologique** active, d'une part pour être au fait de l'état de l'art des technologies et, d'autre part, pour séparer le bon grain de l'ivraie parmi les promesses grandiloquentes d'une industrie qui cherche encore à se définir.

« Dans un tel contexte il y a fort à parier que les compétences mixtes de type métier/technique ou data-science/juridique seront des profils très

recherchés par les entreprises et les cabinets de conseils.»

Dans le même ordre d'idée, il y a fort à parier que la dichotomie classique **MOA/MOE** chère à nos entreprises françaises soit préjudiciable à l'innovation dans les usages de l'IA. Celle-ci est en effet largement antinomique avec ce schéma traditionnel qui voudrait que l'une des parties, la MOE ou le prestataire, supporte l'essentiel du risque associé à l'innovation. Une solide relation de confiance entre partenaires, des intérêts partagés, des risques bien compris et assumés, la mise en œuvre de **méthodes agiles** inspirées des **sciences expérimentales** seront les clés du succès pour les entreprises qui, les premières, défricherons les terra incognita des usages de l'IA.

Pour concrétiser la suite du propos, envisageons à présent une mission d'accompagnement au changement dans le cadre d'un usage bien identifié d'une IA. Supposons à titre d'exemple qu'un service après-vente souhaite déployer un **chatbot** pour **automatiser les réponses aux réclamations des clients** et essayons d'imaginer les questions que soulèverait un tel projet, mené de façon «classique»⁹. Nous distinguerons la phase avant-projet de la phase projet proprement dite.

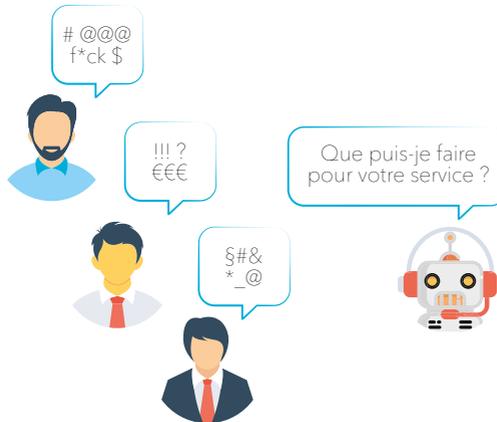


Figure 16 : Automatisation d'un service de réclamation

5.1 AVANT-PROJET

Pour chacune des rubriques ci-dessous nous proposons à titre purement illustratif

⁹ Ceci à des fins de comparaison avec les projets de transformation usuels car une démarche agile serait en réalité plus adaptée à la nature expérimentale d'un tel projet.

une liste de questions qu'un tel projet pourrait susciter.

Etude d'opportunité

L'utilisation d'un chatbot pour répondre à des clients insatisfaits est-elle vraiment de nature à améliorer l'expérience client ou risque-t-elle au contraire de la dégrader? Quelles économies peut-on escompter et à partir quel volume d'appels par jour? Existe-t-il des usages similaires de chatbots dans d'autres organisations sur lesquels il serait possible de capitaliser?

Dossier de choix technologique

Imaginons que le responsable de la veille technologique nous informe qu'il existe actuellement deux types de solutions :

- Des solutions éprouvées et peu onéreuses sous forme des services SaaS qui permettent de répondre aux réclamations récurrentes et simples,
- Des solutions sur mesure et évolutives développées par des startups spécialisées dans la conception de chatbots complexes.

Faut-il **intégrer** les deux **solutions** ou choisir plutôt une solution de type **best of breed**? Comment pourra-t-on monitorer la satisfaction des clients pour améliorer progressivement le service?

Identification de risques

Imaginons que le SAV estime qu'une proportion significative des clients contactent le service de réclamation non pas tant pour obtenir une information mais plutôt pour l'utiliser comme un exutoire émotionnel à leur insatisfaction. La placidité mécanique d'un chatbot ne risque-t-elle pas d'accroître encore l'exaspération de clients qu'il faudrait au contraire rasséréner pour ne pas les perdre? Les mécanismes de reconnaissance vocale sont-ils à même de détecter un état émotionnel qui requiert une réponse formulée par un humain doué d'empathie (même feinte)?

Autre risque: le chatbot sera-t-il en mesure de comprendre qu'un client signal un dysfonctionnement potentiellement dangereux d'un article pour aussitôt en référer à un expert humain?

Etude de faisabilité

Les experts du SAV considèrent qu'il est crucial de pouvoir détecter correctement et rapidement l'état émotionnel d'un client au son de sa voix, ceci pour le réorienter le cas échéant vers un humain compatissant. Aucune étude sérieuse ne répondant pour le moment à cette question il est décidé que le **datalab** du cabinet de conseil construira un prototype dont la seule fonction sera de valider la possibilité d'une telle détection.

Une campagne de sondage sur les appels actuels au SAV permettra par ailleurs de classer les requêtes des clients en trois catégories :

- Les questions récurrentes élémentaires qui pourront être traitées par les services SaaS
- Les questions plus complexes qui requièrent ponctuellement le recours à une IA plus spécialisée et enfin
- Les questions trop complexes qui devront être traitée par un expert humain.

5.2 PROJET

Aide à la conception

Une part importante de la conception consiste sur ce projet imaginaire à trouver les critères qui permettront de distinguer les trois catégories de requêtes mentionnées précédemment.

Accompagnement RH

L'étude de faisabilité a révélé que 80% du personnel du SAV pourra être remplacé par un service de chatbot. La lourde mission des RH sera, en étroite collaboration avec les IRP, d'organiser les **reconversions** du personnel désormais sans affectation.

Organisation et gouvernance

Face à la réduction drastique des effectifs consécutifs à la mise en place du chatbot c'est tout le fonctionnement du SAV qui est à repenser. Dans un premier

temps l'activité du chatbot devra être supervisée étroitement pour éviter les «accidents» et, dans un second temps, pour garantir l'amélioration continue du système. Des formations devront être dispensées aux responsables techniques sur l'optimisation des conversations tenues par le chatbot.



CONCLUSION

Nous nous sommes engagés à ne formuler aucune prophétie et nous tiendrons cet engagement! Une chose cependant apparaît d'ores et déjà clairement: il semble inadéquat d'envisager l'IA comme une simple invention supplémentaire dans une époque déjà fertile en innovations. Les technologies d'IA sont plutôt assimilables à une ressource créatrice de valeur, ou à un socle pour l'innovation future au même titre que le web il y a 25 ans ou même comme l'électricité il y a 150 ans (Ng, 2016). A ce titre le potentiel disruptif de l'IA est probablement est difficile à surestimer.

Chaque technologie ouvre un nouveau champ des possibles, avec ses risques et ses opportunités. Les Grecs anciens, encore eux, avaient déjà formalisés cette idée dans leur concept de **pharmakon** désignant par ce terme l'ambivalence inhérente à toute technologie, qui est tout à la fois un poison et son propre remède. Nulle raison en l'occurrence que l'IA fasse exception. Elle pourra aussi bien contribuer à augmenter notre potentiel d'humanité qu'à la subvertir. A chacun de nous d'utiliser ce qui lui reste de libre arbitre pour faire bon usage de cette extension des possibles sans précédent que pourrait représenter l'IA dans les décennies à venir.

A network diagram consisting of numerous grey circular nodes of varying sizes connected by thin, light grey lines. The nodes are scattered across the page, with a higher density in the upper left and lower right areas. The lines create a complex web of connections between the nodes.

TRAVAUX CITÉS

A FAST LEARNING ALGORITHM FOR DEEP BELIEF NETS [Revue]

Auteurs: Geoffrey E. Hinton Simon Osindero & Yee-Whye Teh

Publication: Journal Neural Computation - Cambridge, MA: MIT Press, juillet 2006 - Vol. 18 issue 7 p. 1527 - 1554

AI MAY PLAY DOCTOR IN THE FUTURE OF HEALTHCARE [En ligne]

Auteur: Furness Dyllan

Publication: Digital Trends - 7/10/2016 - <http://www.digitaltrends.com/cool-tech/artificial-intelligence-chatbots-are-revolutionizing-healthcare/>

AI'S LANGUAGE PROBLEM [Article]

Auteur: Knight Will

Publication: MIT Technology Review, 2016 - <https://www.technologyreview.com/s/602094/ai-language-problem/>

ARTIFICIAL INTELLIGENCE: A MODERN APPROACH (3RD EDITION) [Ouvrage]

Auteurs: Norvig Peter & Russel Stuart J.

Publication: Pearson, 2009 - p. 1152

BARACK OBAMA, NEURAL NETS, SELF-DRIVING CARS, AND THE FUTURE OF THE WORLD [En ligne]

Auteur: Dadich Scott

Publication: Wired - Final Frontier - 24/8/2016 - <https://www.wired.com/2016/10/president-obama-mit-joi-ito-interview/>

BUILDING HIGH-LEVEL FEATURES USING LARGE SCALE UNSUPERVISED LEARNING [Conférence]

Auteurs: Quoc V. Le Marc'Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg S. Corrado, Jeffrey Dean & Andrew Y. Ng

Publication: éd. Learning ICML 2012: 29th International Conference on Machine - Edinburgh : [s.n.], 2012

DEEP VISUAL-SEMANTIC ALIGNMENTS FOR GENERATING IMAGE DESCRIPTIONS [En ligne]

Auteur: Andrej Karpathy Li Fei-Fei

Publication: 2015. - <http://cs.stanford.edu/people/karpathy/deepimagesent/>

HOW ARTIFICIAL INTELLIGENCE WILL REDEFINE MANAGEMENT [En ligne]

Auteurs: Vegard Kolbjørnsrud Richard Amico & Robert J. Thomas

Publication: Harvard Business Review - 02 11 2016 - <https://hbr.org/2016/11/how-artificial-intelligence-will-redefine-management>

HUMAN-LEVEL CONTROL THROUGH DEEP REINFORCEMENT LEARNING [Revue]

Auteurs: Volodymyr Mnih Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Demis Hassabis

Publication: Nature - février 2015 - Vol. 518 - pp. 529-533

PERCEPTONS: AN INTRODUCTION TO COMPUTATIONAL GEOMETRY [Ouvrage]

Auteurs: Papert Marvin Minsky & Seymour

Publication: Cambridge: The MIT Press, 1972.

THE ETHICS OF ARTIFICIAL INTELLIGENCE [En ligne]

Auteur: Loukides Mike

Publication: O'Reilly Ideas - 14/11/2016 - <https://www.oreilly.com/ideas/the-ethics-of-artificial-intelligence>

WEAPONS OF MATH DESTRUCTION [Ouvrage]

Auteur: O'Neil Cathy

Publication: Crown , 2016 - p. 272

WHAT ARTIFICIAL INTELLIGENCE CAN AND CAN'T DO RIGHT NOW [En ligne]

Auteur: Ng Andrew

Publication: HBR - 09/11/2016 - https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now?imm_mid=0ea9bf&cmp=em-data-na-na-newsltr_ai_20161114

WHAT IS ARTIFICIAL INTELLIGENCE ? [Rapport]

Auteurs: Loukides Mike & Lorica Ben

Publication: O'Reilly, 2016 - <https://www.oreilly.com/ideas/what-is-artificial-intelligence>

WHY ARTIFICIAL INTELLIGENCE IS THE NEW ELECTRICITY [En ligne]

Auteur: Ng Andrew

Publication: Inc.Video - 2016 - <http://www.inc.com/andrew-ng/why-artificial-intelligence-is-the-new-electricity.html>

CONCEPTION :

Auteurs: Pirmin Lemberger - [Data Scientist & Directeur DataLab Weave](#)
Jonathan Lapan - [Senior Manager](#)
Olivier Reisse - [Fondateur de weave Business Technology](#)

Designer: Pierre Wustmann - [Designer](#)

©2017 par Weave Business Technology publié et distribué en Français dans le monde

Première publication weave.eu en 2017.



37 rue du rocher
75008 Paris, France

Tél.: +33 (0)1 58 44 58 44

Email: contact@weave.eu

Site web: weave.eu



37 rue du rocher
75008 Paris, France

Tel. : +33 (0)1 58 44 58 44
weave.eu