

N° 4594

N° 464

ASSEMBLÉE NATIONALE

SÉNAT

CONSTITUTION DU 4 OCTOBRE 1958
QUATORZIÈME LÉGISLATURE

SESSION ORDINAIRE 2016 - 2017

Enregistré à la présidence de l'Assemblée
nationale

Enregistré à la présidence du Sénat

le 15 mars 2017

le 15 mars 2017

RAPPORT

au nom de

**L'OFFICE PARLEMENTAIRE D'ÉVALUATION
DES CHOIX SCIENTIFIQUES ET TECHNOLOGIQUES**

**POUR UNE INTELLIGENCE ARTIFICIELLE
MAÎTRISÉE, UTILE ET DÉMYSTIFIÉE**

PAR

M. Claude de GANAY, député, et Mme Dominique GILLOT, sénatrice

Tome II : Annexes

Déposé sur le Bureau de l'Assemblée nationale
par M. Jean-Yves LE DÉAUT,
Président de l'Office

Déposé sur le Bureau du Sénat
par M. Bruno SIDO,
Premier vice-président de l'Office

Composition de l'Office parlementaire d'évaluation des choix scientifiques et technologiques

Président

M. Jean-Yves LE DÉAUT, député

Premier vice-président

M. Bruno SIDO, sénateur

Vice-présidents

M. Christian BATAILLE, député
Mme Anne-Yvonne LE DAIN, députée
M. Jean-Sébastien VIALATTE, député

M. Roland COURTEAU, sénateur
M. Christian NAMY, sénateur
Mme Catherine PROCACCIA, sénateur

DÉPUTÉS

M. Bernard ACCOYER
M. Gérard BAPT
M. Alain CLAEYS
M. Claude de GANAY
Mme Françoise GUÉGOT
M. Patrick HETZEL
M. Laurent KALINOWSKI
M. Alain MARTY
M. Philippe NAUCHE
Mme Maud OLIVIER
Mme Dominique ORLIAC
M. Bertrand PANCHER
M. Jean-Louis TOURAINE

SÉNATEURS

M. Patrick ABATE
M. Gilbert BARBIER
Mme Delphine BATAILLE
M. Michel BERSON
M. François COMMEINHES
Mme Catherine GÉNISSON
Mme Dominique GILLOT
M. Alain HOUPERT
Mme Fabienne KELLER
M. Jean-Pierre LELEUX
M. Gérard LONGUET
M. Pierre MÉDEVIELLE
M. Franck MONTAUGÉ
M. Hervé POHER

SOMMAIRE

	<u>Pages</u>
ANNEXE 1 : LISTE DES DÉPLACEMENTS DES RAPPORTEURS	7
ANNEXE 2 : PANORAMAS DE L'INTELLIGENCE ARTIFICIELLE	9
1. <i>Aux États-Unis</i>	9
2. <i>En Chine</i>	21
3. <i>Au Japon</i>	27
4. <i>Au Royaume-Uni</i>	36
5. <i>En Suisse</i>	43
ANNEXE 3 : SUR LE DROIT COMPARÉ DE LA ROBOTIQUE	49
ANNEXE 4 : CONTRIBUTION DE M. RAJA CHATILA, DIRECTEUR DE L'INSTITUT DES SYSTÈMES INTELLIGENTS ET DE ROBOTIQUE (ISIR) « ROBOTIQUE ET INTELLIGENCE ARTIFICIELLE »	75
ANNEXE 5 : CONTRIBUTION DE M. PATRICK ALBERT, ENTREPRENEUR ET CHERCHEUR : RELEVER DÉMOCRATIQUEMENT LES DÉFIS DE L'INTELLIGENCE ARTIFICIELLE	79
ANNEXE 6 : RÉPONSES ÉCRITES AUX RAPPORTEURS	95
I. RÉPONSES D'AXELLE LEMAIRE, SECRÉTAIRE D'ÉTAT CHARGÉE DU NUMÉRIQUE ET DE L'INNOVATION, ET DE THIERRY MANDON, SECRÉTAIRE D'ÉTAT CHARGÉ DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE, AU QUESTIONNAIRE DE VOS RAPPORTEURS	95
II. RÉPONSES À LA CONSULTATION ORGANISÉE PAR LE PARIS MACHINE LEARNING MEETUP EN VUE DU RAPPORT DE L'OPECST	108
ANNEXE 7 : COMPTE RENDU DE L'AUDITION PUBLIQUE DU 19 JANVIER 2017	115
I. INTRODUCTION	115
1. <i>M. Bruno Sido, sénateur, premier vice-président de l'OPECST</i>	115
2. <i>M. Jean-Yves Le Déaut, député, président de l'OPECST</i>	116
II. PREMIÈRE TABLE RONDE, PRÉSIDÉE PAR M. CLAUDE DE GANAY, RAPPORTEUR : LES TECHNOLOGIES RELEVANT DE L'INTELLIGENCE ARTIFICIELLE	119
1. <i>Mme Dominique Gillot, sénatrice, membre de l'OPECST, rapporteure</i>	119
2. <i>M. Claude de Ganay, député, membre de l'OPECST, rapporteur</i>	119
3. <i>M. Jean-Gabriel Ganascia, professeur à l'Université Pierre-et-Marie-Curie, Paris-VI</i>	120
4. <i>M. Gérard Sabah, directeur de recherche honoraire au CNRS</i>	123
5. <i>M. Yves Demazeau, président de l'Association française pour l'intelligence artificielle</i>	127
6. <i>M. Bertrand Braunschweig, directeur du centre Inria de Saclay</i>	134
7. <i>M. David Sadek, directeur de la recherche à l'Institut Mines-Télécom</i>	138
8. <i>M. Jean-Daniel Kant, maître de conférences à l'Université Pierre-et-Marie-Curie Paris_VI</i>	140

9. M. Benoît Le Blanc, directeur adjoint de l'École nationale supérieure de cognitive	142
10. Débat.....	145
III. DEUXIÈME TABLE RONDE, PRÉSIDIÉE PAR MME DOMINIQUE GILLOT, RAPPORTEURE : DIMENSIONS STRATÉGIQUES EN MATIÈRE DE RECHERCHE EN INTELLIGENCE ARTIFICIELLE	150
1. M. Jean-Marc Merriaux, directeur général de Canopé	150
2. M. François Taddei, directeur du Centre de recherche interdisciplinaire (CRI)	152
3. M. Olivier Esper, responsable des affaires publiques de Google France	154
4. Mme Delphine Reyre, directrice Europe des affaires publiques de Facebook	156
5. M. Laurent Massoulié, directeur du Centre de recherche commun Inria-Microsoft	158
6. M. Dominique Cardon, professeur de sociologie à l'Institut d'Études Politiques de Paris/Médialab	159
7. M. Gilles Babinet, entrepreneur, digital champion auprès de la Commission européenne	161
8. Débat.....	163
IV. TROISIÈME TABLE RONDE, PRÉSIDIÉE PAR MME DOMINIQUE GILLOT, RAPPORTEURE : QUESTIONS POLITIQUES, SOCIÉTALES ET ÉCONOMIQUES LIÉES À L'IRRUPTION DES TECHNOLOGIES DE L'INTELLIGENCE ARTIFICIELLE	172
1. M. Henri Verdier, directeur interministériel du numérique	172
2. Mme Marie-Claire Carrère-Gée, présidente du Conseil d'orientation pour l'emploi (COE)	174
3. M. Laurent Alexandre, entrepreneur (DNA vision)	177
4. M. Jean-Christophe Baillie, entrepreneur (Novaquark)	182
5. M. Jean-Claude Heudin, directeur de l'Institut de l'Internet et du multimédia	185
6. Débat.....	187
V. INTERVENTION DE LA MINISTRE : AXELLE LEMAIRE, SECRÉTAIRE D'ÉTAT CHARGÉE DU NUMÉRIQUE ET DE L'INNOVATION	192
Débat	195
VI. QUATRIÈME TABLE RONDE, PRÉSIDIÉE PAR M. CLAUDE DE GANAY, RAPPORTEUR : ENJEUX ÉTHIQUES DE L'INTELLIGENCE ARTIFICIELLE	198
1. M. Gilles Dowek, directeur de recherche Inria, professeur attaché à l'ENS Paris-Saclay	198
2. Mme Laurence Devillers, professeur à l'Université Paris-Sorbonne/LMSI-CNRS	201
3. M. Serge Abiteboul, directeur de recherche Inria	203
4. M. Jean Ponce, professeur et directeur du département d'informatique de l'École normale supérieure (ENS).....	205
5. M. Serge Tisseron, psychiatre, chercheur associé à l'Université Paris Diderot-Paris VII.....	206
VII. CINQUIÈME TABLE RONDE, PRÉSIDIÉE PAR MME DOMINIQUE GILLOT, RAPPORTEURE : DÉFIS JURIDIQUES INHÉRENTS AUX USAGES DE L'INTELLIGENCE ARTIFICIELLE	209
1. Intervention de Mady Delvaux, députée européenne (Luxembourg - groupe S&D), rapporteuse du groupe de travail sur la robotique et l'intelligence artificielle	209
2. Mme Isabelle Falque-Pierrotin, présidente de la CNIL	212
3. M. Rand Hindi, membre du Conseil national du numérique, pilote du groupe de travail sur l'intelligence artificielle, président de SNIPS	213
4. M. Olivier Guilhem, directeur juridique chez SoftBank Robotics (ex Aldebaran)	214
5. M ^e Alain Bensoussan, avocat, président de l'Association du droit des robots	216
6. Débat.....	217
VIII. CONCLUSION.....	218

ANNEXE 8 : COMPTES RENDUS DES AUDITIONS BILATÉRALES CONDUITES PAR LES RAPPORTEURS	219
I. AUDITIONS DU 25 OCTOBRE 2016	219
1. <i>M. Stéphane Mallat, professeur à l'École normale supérieure (ENS), chercheur en mathématiques appliquées.....</i>	219
2. <i>M. Patrick Albert, entrepreneur (créateur de ILOG), chercheur et pionnier dans le domaine de l'intelligence artificielle</i>	221
II. AUDITIONS DU 8 NOVEMBRE 2016	222
1. <i>M. Marc Mézard, directeur de l'École normale supérieure (ENS).....</i>	222
2. <i>M. Raja Chatila, directeur de recherche au Centre national de la recherche scientifique (CNRS), directeur de l'Institut des systèmes intelligents et de robotique (ISIR)</i>	224
III. AUDITIONS DU 9 NOVEMBRE 2016.....	226
1. <i>M^e Alain Bensoussan, avocat, président de l'association pour les droits des robots et M^e Marie Soulez, avocate spécialisée sur les TIC dans son cabinet</i>	226
2. <i>M. Henri Verdier, directeur interministériel du numérique, ancien entrepreneur et spécialiste du numérique.....</i>	229
3. <i>M. Laurent Alexandre, président de DNA Vision, fondateur de Doctissimo, chirurgien- urologue</i>	230
4. <i>M. Pierre-Yves Oudeyer, directeur de recherche Inria, directeur du laboratoire Flowers, président du comité technique des systèmes cognitifs et développementaux de l'IEEE (Institut des ingénieurs électriciens et électroniciens)</i>	232
IV. AUDITIONS DU 24 NOVEMBRE 2016.....	240
1. <i>Mme Flora Fischer, chargée de programme de recherche au CIGREF, Club informatique des grandes entreprises françaises, pilote du groupe de travail sur l'intelligence artificielle en entreprise</i>	240
2. <i>M. Max Dauchet, professeur émérite à l'Université de Lille, président de la Commission de Réflexion sur l'Éthique de la Recherche en sciences et technologies du numérique (CERNA) d'Allistene, alliance des sciences et technologies du numérique</i>	243
3. <i>M. Cédric Sauviat, ingénieur, président de l'association française contre l'intelligence artificielle (AFCIA) et Mme Marie David, ingénieur, éditrice, membre du bureau de l'association.....</i>	246
V. AUDITIONS DU 28 NOVEMBRE 2016	257
1. <i>M. Claude Berrou, professeur à Télécom Bretagne (Institut Mines-Télécom), chercheur en électronique et informatique, membre de l'Académie des sciences</i>	257
2. <i>M. Nicolas Cointe et Mme Fiona Berreby, chercheurs en thèse de doctorat sur l'éthique de l'intelligence artificielle</i>	259
3. <i>Mme Laurence Devillers, professeur d'informatique à l'Université Paris-Sorbonne et directrice de recherche du CNRS au Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (Limsi de Saclay).....</i>	262
VI. AUDITIONS DU 30 NOVEMBRE 2016.....	264
1. <i>M. David Sadek, directeur de la recherche de l'Institut Mines-Télécom, spécialiste en intelligence artificielle</i>	264
2. <i>M. Dominique Sciamma, directeur de l'école de design « Strate » à Sèvres</i>	266
3. <i>M. François Taddéi, directeur du Centre de recherches interdisciplinaires (Inserm, Université Paris-Descartes), biologiste</i>	269
4. <i>M. Igor Carron, entrepreneur, organisateur du principal « meet-up » en intelligence artificielle en France intitulé « Paris Machine Learning »</i>	270
5. <i>M. Jill-Jënn Vie, chercheur en thèse de doctorat à l'École normale supérieure Paris- Saclay.....</i>	271



ANNEXE 1 : LISTE DES DÉPLACEMENTS DES RAPPORTEURS

Vos rapporteurs ont effectué les déplacements suivants :

- un déplacement aux **États-Unis d'Amérique** du 22 au 29 janvier 2017, pour rencontrer des spécialistes de l'intelligence artificielle, à Washington à l'Institut de technologie du Massachusetts (MIT), à la *Harvard Kennedy School of Government*, à l'Université de Washington, à l'Université de Stanford, à l'Université de Berkeley, ainsi que des représentants de Facebook (Menlo Park, Palo Alto), de Google (Mountain View), d'Apple (Cupertino) et de Salesforce (San Francisco)... ;

- trois déplacements en **Europe** : à **Genève** du 21 au 22 septembre 2016 (HBP et BBP du *Brain Mind Institute* de l'École polytechnique fédérale de Lausanne, EPFL), au **Royaume-Uni** du 13 au 16 décembre 2016 (Chambre des Communes, Royal Society, *Alan Turing Institute*, *Future of humanity Institute* de l'Université d'Oxford, CSER et LCFI de Cambridge...) et à **Bruxelles** du 8 au 9 février 2017 (laboratoire d'intelligence artificielle de l'Université libre de Bruxelles, institutions européennes...) ;

- un déplacement en France métropolitaine, à **Arcachon** du 26 au 30 septembre 2016, pour participer à un séminaire sur l'éthique de l'intelligence artificielle organisé par la commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene (CERNA).

ANNEXE 2 : PANORAMAS DE L'INTELLIGENCE ARTIFICIELLE

1. Aux États-Unis

Note réalisée avec le concours du service scientifique de l'Ambassade de France aux États-Unis et en particulier du Consulat de San Francisco

S'il fallait un indice de la prise de conscience collective aux États-Unis de l'importance critique du développement de l'intelligence artificielle dans ses perspectives scientifiques et technologiques mais également et peut-être surtout économiques et sociétales, le fait que le Président Barack Obama ait choisi de consacrer la dernière de ses grandes apparitions publiques consacrées à la science et à la technologie à un entretien extrêmement médiatisé avec Joi Ito, le directeur du prestigieux MIT Media Lab, publié dans le plus grand magazine de vulgarisation technologique du monde, et entièrement dédié à une discussion autour de l'intelligence artificielle¹ constitue probablement un assez bon indicateur de ce que cette thématique technologique est le sujet brûlant du moment aux États-Unis, que ce soit dans l'expression de politiques publiques, la consolidation d'une puissance de recherche inégalée ou le développement rapide et sans précédent d'activités économiques liées.

1. Les politiques publiques liées au développement de l'intelligence artificielle

Consciente de la tension entre l'impact potentiellement bénéfique des progrès technologiques dans le champ de l'intelligence artificielle et de leur percolation dans un nombre croissant de champs d'activité économique d'une part, et de ses potentielles implications structurantes sur la société, relayées publiquement de surcroît par diverses personnalités en vue du monde scientifique et industriel, d'autre part, la Maison Blanche a mis en place en mai 2016 un sous-comité spécifique au sein du *National Science and Technology Council* (NSTC), chargé de suivre les évolutions du secteur et de coordonner les activités fédérales sur le sujet. En parallèle ont été tenues quatre sessions de travail publiques entre les mois de mai et juillet 2016, visant à engager la discussion avec le grand public et surtout à produire une large évaluation des opportunités, risques, et implications réglementaires et sociales de l'intelligence artificielle, de même qu'une série de recommandations dotées d'un plan stratégique afin de se donner les moyens de les mettre en œuvre. Les conclusions de ces travaux ont donné lieu à pas

¹ "Barack Obama, neural nets, self-driving cars and the future of the world", *entretien du 24 août 2016 dans WIRED*. <https://www.wired.com/2016/10/president-obama-mit-joi-ito-interview/>

moins de trois rapports produits par l'administration Obama entre octobre et décembre 2016.

1.1 *Preparing the Future of Artificial Intelligence* (NSTC, octobre 2016)

Ce premier rapport, assez peu détaillé, a pour objectif de produire un cadre général à la réflexion nationale, en amorçant un état des lieux de la recherche et des applications actuelles tout en posant des jalons initiaux quant à la possibilité et la nature d'une régulation.

En particulier, le rapport conclut, avec un optimisme affiché, que l'intelligence artificielle et l'apprentissage automatique font émerger le potentiel d'améliorer la vie des citoyens en permettant de résoudre certains des plus grands enjeux sociétaux. La santé, les transports, l'environnement, la justice ou encore l'efficacité du gouvernement lui-même sont autant de secteurs applicatifs mis en avant.

Cela étant, les auteurs insistent sur l'importance de la réglementation dans l'accompagnement de ces avancées technologiques. Si un consensus se dégage pour conclure qu'une réglementation générale de la recherche en IA semble inapplicable à l'heure actuelle, et, partant, la réglementation actuelle est pour l'heure suffisante, l'approche explicite guidant les prochains arbitrages se devra de procéder selon le principe suivant : évaluer les risques que l'implémentation de l'IA pourrait réduire, de même que ceux qu'elle pourrait augmenter. L'analyse comparée des risques et des bénéfices permettra de justifier les futures décisions d'ordre réglementaire. En outre, le rapport insiste sur l'importance d'ajuster les réglementations afin de réduire, au maximum, les coûts et les barrières à l'innovation sans mettre en danger la sécurité du public ou la compétition équitable sur le marché.

Ce n'est pas pour autant diminuer le rôle pivot des autorités fédérales. De ce point de vue, le Gouvernement a plusieurs rôles à jouer : un rôle d'organisateur du débat public et d'arbitre des mesures à mettre en place à l'échelle du pays ; un rôle de suivi attentif de la sécurité et de la neutralité des applications développées ; un rôle de producteur de réglementations encourageant l'innovation tout en protégeant le public ; un rôle d'accompagnateur de la diffusion de ces technologies tout en protégeant certains secteurs au besoin afin d'éviter des contrecoups économiques dévastateurs ; un rôle de soutien et de financeur de projets de recherche faisant avancer le domaine ; et enfin un rôle d'adoption en son sein même de ces avancées afin d'assurer un service public de meilleure qualité.

En particulier, dans cette position, le gouvernement fédéral a un rôle clé à jouer dans l'avancement de l'IA en termes de R&D, en particulier à travers la production d'une main-d'œuvre en nombre suffisant, ainsi que d'un haut niveau de qualification et de diversité technique, non seulement

d'un point de vue professionnel, mais également, et de manière plus large, du point de vue de la formation générale de la population.

En termes d'impact économique, le rapport conclut que le premier effet à court terme sera celui de l'automatisation d'un nombre grandissant de tâches. Si la productivité générale et la création de richesse ont toutes les chances d'être positivement impactées, l'impact sur l'emploi est lui plus difficile à évaluer. Ce dernier sera inégal en fonction des professions, et le sous-comité suggère que les professions à faible salaire seront les plus touchées, et que l'automatisation aggrave le fossé entre travailleurs hautement formés et travailleurs faiblement éduqués, s'accompagnant potentiellement d'une augmentation des inégalités économiques. Dans cette optique, le rôle du gouvernement serait ici d'assurer le maintien de certaines catégories de travailleurs pouvant être considérés comme complémentaires aux machines automatisées, plutôt que concurrents. En sus, les politiques publiques pourront s'efforcer d'assurer un partage général des bénéfices économiques engendrés.

1.2 The National Artificial Intelligence Research & Development Strategic Plan (NSTC, octobre 2016)

Ce plan stratégique, publié en octobre 2016, présente une série d'objectifs à poursuivre pour la recherche en IA financée par des fonds publics fédéraux, avec l'ambition de couvrir à la fois les efforts de recherche directement produits par des entités fédérales et ceux des organisations externes, au premier rang desquelles les universités, tout en essayant de mettre l'accent sur les domaines dans lesquelles l'industrie est moins susceptible d'intervenir. Sept priorités en constituent le cœur stratégique :

- Investissement fédéral soutenu dans la recherche à long terme afin de produire des percées scientifiques et technologiques dans les dix prochaines années, en particulier dans les nouvelles méthodologies nécessaires à la découverte de savoir à l'heure des grands volumes de données, l'amélioration des capacités de perception des systèmes d'intelligence artificielle, la compréhension profonde des capacités théoriques et des limites propres au développement de l'intelligence artificielle, la poursuite des efforts visant au développement d'une IA générique par opposition aux intelligences artificielles spécifiques, la nécessité de développer des systèmes d'intelligence artificielle extensible, l'accélération de la recherche sur une intelligence artificielle s'approchant des capacités humaines.

- Développement de méthodes de collaboration entre homme et intelligence artificielle : systèmes fonctionnant en parallèle, en substitution au moment où l'individu atteint ses limites cognitives ou en substitution totale.

- Compréhension accrue des implications légales, sociétales et éthiques, en particulier développement de méthodes permettant de concevoir des systèmes d'intelligence artificielle conformes aux principes éthiques, légaux et sociétaux des États-Unis. Plus précisément, le rapport insiste sur l'importance d'assurer la justice, la transparence et la responsabilité des systèmes, dès la phase de conception.
- Sécurité et sûreté des systèmes, en particulier dans l'adaptation à des environnements complexes et/ou incertains.
- Développement de bases de données publiques partagées et d'environnements pour l'entraînement et le test de systèmes d'intelligence artificielle, à commencer par la mise à disposition de jeux de données fédéraux existants.
- Développement d'un spectre large de standards visant à assurer que les technologies émergentes répondent à des objectifs précis en termes de fonctionnalité et d'interopérabilité, ainsi qu'en termes de sécurité et de performance.
- Évaluation des besoins nationaux en termes de main-d'œuvre.

1.3 Artificial Intelligence, Automation and the Economy (Executive Office of the President, décembre 2016)

Dans la foulée des deux précédents, l'*Executive Office of the President* vient de produire un troisième rapport, cosigné par plusieurs branches, et centré sur les implications économiques, en particulier sur le marché du travail, du développement de l'intelligence artificielle.

S'il en souligne comme toujours les bénéfices économiques attendus, il soulève la question de la répartition de ces bénéfices, en évoquant en particulier le risque que cette répartition asymétrique contribue de manière décisive à creuser les inégalités entre les travailleurs hautement qualifiés et les autres. Bien qu'il soit hasardeux d'essayer de prédire exactement quels emplois seront le plus tôt impactés par l'automatisation conduite par l'intelligence artificielle (aussi bien en termes de destruction d'emplois que de création de nouveaux emplois, par exemple de supervision de l'intelligence artificielle), le rapport pointe que, d'après les experts, l'ampleur des volumes d'emplois directement menacés dans les décennies à venir serait comprise entre 9 % et 47 % du volume total d'emplois, le plus largement concentrés au sein des groupes de travailleurs les moins diplômés et les moins bien rémunérés, ce qui implique qu'un des effets directs de cette vague d'automatisation pour ce groupe résidera dans la pression vers le bas des salaires. Le rapport va même jusqu'à pointer très explicitement la menace selon laquelle l'automatisation en question pourrait bien ne bénéficier qu'à quelques-uns, de par la nature hautement hégémonique (« *winner-takes-most* ») des marchés liés aux technologies de l'information.

Cela étant, le rapport insiste par contraste sur la capacité des politiques publiques et des incitations institutionnelles à réguler les effets du changement technologique (« *Technology is not destiny* »). Il formule dès lors un certain nombre de propositions stratégiques dont il estime qu'il est encore plus important qu'elles soient mises en œuvre au regard du contexte d'émergence massive à moyen terme de l'intelligence artificielle et de l'automatisation :

- Investir dans et développer l'intelligence artificielle en reconnaissant les bénéfices, à la fois en termes de croissance de la productivité et de volonté de rester à la pointe de l'économie de l'innovation, en menant des politiques volontaristes d'accès à cette économie, en particulier du point de vue de la diversité.
- Éduquer et former les Américains aux emplois du futur, que ce soit de manière précoce, dans la capacité des politiques d'enseignement à rendre accessible financièrement l'enseignement postsecondaire, ou dans l'apprentissage tout au long de la carrière professionnelle.
- Soutenir les travailleurs en transition (en particulier en accroissant la disponibilité des dispositifs de sécurisation de l'emploi et de la transition vers l'emploi, mais également l'accroissement des salaires).

2. Quelques repères sur la recherche en intelligence artificielle aux États-Unis

L'objectif de cette section est de proposer quelques repères sur la recherche en intelligence artificielle aux États-Unis, sans volonté d'exhaustivité dans la présentation de ce paysage scientifique.

2.1 Quelques grands centres

L'excellence de certaines universités comme le MIT, Stanford et Harvard n'est plus à démontrer et celles-ci abritent des équipes et des chercheurs qui jouent un rôle important dans la recherche en IA.

MIT : Computer Science and Artificial Intelligence Laboratory (CSAIL)

Le Massachusetts Institute of Technology (MIT) possède avec le CSAIL un centre dédié à l'IA qui est remarquable de par la forte diversité dans les thématiques traitées, le nombre et la qualité de ses équipes de recherche. L'ensemble des cinquante équipes regroupe mille deux cents personnes dont environ trois cents chercheurs, huit cents étudiants (six cents *graduate students*, deux-cents *undergraduate students*). Par rapport à la démographie usuelle en France, il est intéressant de noter la proportion beaucoup plus forte d'étudiants (pour ces thématiques, la répartition usuelle en France est plutôt de 50 % de chercheurs et 50 % d'étudiants).

Carnegie Mellon University (CMU)

Carnegie Mellon University (CMU) est une université moins connue mais pourtant essentielle dans le paysage de l'informatique avec ses douze récipiendaires du prix Turing (par comparaison, la France n'en compte qu'un seul), qui est l'équivalent du prix Nobel pour l'informatique. Avec son institut de Robotique, qui a été le premier créé aux États-Unis en 1970, CMU a joué et joue toujours un rôle central dans ce domaine. Cet institut abrite notamment le *Field Robotic Center* qui conçoit des robots capables de travailler dans des conditions extrêmes (explorations polaires et explorations spatiales par exemple) et le *NanoRobotics lab* dont l'objet d'étude est centré sur les nanorobots avec notamment des applications en médecine très prometteuses.

Stanford Artificial Intelligence Laboratory (SAIL)

Pour illustrer la forte interaction entre le monde académique et le monde industriel, on peut noter qu'en liaison avec le SAIL, l'Université de Stanford a créé un laboratoire conjoint avec Toyota, le *Stanford-Toyota Research Center*, dont l'objet d'étude n'est pas exclusivement le véhicule autonome mais, de manière plus large, le développement des interactions entre l'homme et les machines intelligentes. Ce laboratoire pluridisciplinaire regroupe vingt et un chefs de projets qui travaillent sur douze thématiques liées à l'intelligence artificielle¹.

Berkeley Artificial Intelligence Research (BAIR)

Beaucoup plus petit que le CSAIL du MIT (vingt-cinq professeurs pour une centaine d'étudiants diplômés), ce laboratoire couvre les principaux domaines de l'IA (vision par ordinateur, apprentissage automatique, traitement automatique des langues et robotique). Ce laboratoire se caractérise par l'excellence scientifique de certains de ces membres dont quelques professeurs emblématiques comme Michael I. Jordan et Stuart Russell, mais aussi le très prometteur Sergey Levine qui applique avec succès des techniques d'apprentissage automatique en robotique². On notera que l'Université de Californie à Berkeley vient de lancer un Centre de Recherche dédié aux interactions entre l'homme et l'intelligence artificielle (*Centre for Human-Compatible Artificial Intelligence - CHCAI*) piloté par Stuart Russell.

2.2 Quelques acteurs importants aux États-Unis

Les États-Unis ont une capacité d'attractivité exceptionnelle et ont réussi à attirer plusieurs grands scientifiques de la discipline.

¹ <http://aicenter.stanford.edu/research/#people>

² <https://www.technologyreview.com/lists/innovators-under-35/2016/pioneer/sergey-levine/>

2.2.1 Les pionniers

Alan Turing (1912-1954) : d'origine britannique, il est considéré comme le père fondateur de l'informatique avec ses contributions fondamentales sur la décidabilité. Il est resté célèbre dans le monde de l'intelligence artificielle par sa proposition du test de Turing qui vise à déterminer si une machine est dotée d'intelligence artificielle. Ce test qui consiste à tester la capacité de confondre un ordinateur avec un interlocuteur humain est aujourd'hui remis en cause pour prendre en compte la grande variété des critères et caractéristiques d'un être ou d'une machine intelligente.

John McCarthy (1927-2011) et Marvin Lee Minsky (1927-2016) : John McCarthy a été professeur à Stanford et au MIT. Ce mathématicien est à l'origine du terme « intelligence artificielle » et l'inventeur du langage de programmation LISP qui était très populaire du temps des systèmes experts. Ses travaux autour de la logique symbolique sont essentiels en théorie des jeux (il est à l'origine de l'élagage alpha-bêta permettant de réduire l'espace d'exploration dans un arbre de possibilités). Il est avec Marvin Minsky l'un des organisateurs de l'école d'été à Dartmouth où le concept d'IA a émergé. Marvin Minsky est un spécialiste des sciences cognitives, co-fondateur du laboratoire d'intelligence artificielle au MIT. Il fut notamment le créateur du premier calculateur basé sur les réseaux de neurones.

2.2.2 Les contemporains

Les trois prix Turing américains en IA : Edward Feigenbaum (1994), Leslie Valiant (2010) et Judea Pearl (2011) : Après des études à l'Université Carnegie-Mellon, Edward Feigenbaum est devenu professeur d'informatique et codirecteur scientifique du *Knowledge Systems Laboratory* à l'Université de Stanford. Il a obtenu le Prix Turing pour ses travaux dans les domaines de l'étude et la construction de systèmes d'intelligence artificielle à grande échelle. Judea Pearl était professeur à New York University (NYU) et a reçu le prix Turing pour ses travaux sur les calculs de probabilités et les approches bayésiennes en Intelligence Artificielle. Leslie Valiant est un britannique, professeur à Harvard, et reconnu notamment pour son modèle PAC en théorie de l'apprentissage.

Les récipiendaires de IJCAI/AAAI (principales conférences du domaine), Michael I. Jordan et Barbara J. Grosz : Michael I. Jordan est professeur à Berkeley et lauréat du prestigieux prix IJCAI en 2016 pour ses travaux en apprentissage automatique. Michael I. Jordan est également membre de l'Académie des sciences et a obtenu la chaire d'excellence de la Fondation des Sciences Mathématiques de Paris en 2012. Barbara J. Grosz est professeur à l'Université d'Harvard et spécialiste du Traitement Automatique des langues et des systèmes multi-agents. Elle a reçu le prix IJCAI en 2015 et le prix de l'ACM/AAAI Allen Newell en 2009.

Une responsable de la société savante, Manuela Veloso : Manuela Veloso est professeur à CMU et préside la principale société savante en IA, *l'Association for the Advancement of Artificial Intelligence (AAAI)*. Cette scientifique qui travaille dans le domaine des multi-agents défend une approche baptisée « *Symbiotic Autonomy* » qui est basée sur la collaboration entre humains et robots en fonction du contexte.

Des scientifiques impliqués dans les enjeux sociétaux liés à l'IA, de Gary Marcus à Stuart Russell, en passant par Elon Musk : Gary Marcus est professeur de psychologie à New York University et est connu pour ses romans et aussi son leadership dans le mouvement pour remplacer le test de Turing par une série de tests baptisée « *Turing Olympics* »¹. On ne présente plus Elon Musk : ce Sud-Africain d'origine est devenu créateur d'entreprises innovantes (Tesla Motors, SpaceX) et symbolise parfaitement la puissance d'innovation technologique américaine. Il est le cofondateur et le coprésident d'OpenAI² qui vise à promouvoir *l'open source* en intelligence artificielle. La promotion du logiciel libre est particulièrement importante pour la recherche mais aussi pour l'émergence de start-up en particulier en France dans le domaine. Quant à Stuart Russell, professeur à Berkeley, est un scientifique de renom qui explore différentes facettes de l'intelligence dans ses travaux. Il est notamment connu pour ses contributions sur l'apprentissage par renforcement hiérarchique et ses réflexions sur les impacts sociétaux de l'IA, notamment le danger potentiel des armes autonomes. Stuart Russell a été titulaire d'une chaire Blaise Pascal de la Région Île-de-France de 2012 à 2014.

2.3 Quelques tendances dans différents domaines de l'intelligence artificielle

Vision par ordinateur (*Computer Vision*) : La principale tendance dans ce domaine de recherche est caractérisée par le passage de méthodes dites supervisées (nécessitant une intervention humaine) à des méthodes peu supervisées voire non supervisées. L'objectif est de pouvoir indexer des quantités de données de plus en plus grandes. L'indexation consiste à proposer des algorithmes pour extraire des éléments sémantiques caractéristiques liés à la perception telle que la reconnaissance d'objets, la détection de séquences dans des vidéos, la description de contenus (personnes, activités, mouvements, ...). Ce domaine est essentiel pour faire face aux défis scientifiques posés par le *big data*.

Apprentissage automatique (*Machine Learning*) : Ici encore, comme le souligne Yann LeCun, la principale tendance est le développement de l'apprentissage non supervisé. Une autre tendance forte est la recherche sur la compréhension par des non-experts des conclusions, actions et décisions retournées par les programmes d'IA. La DARPA vient de lancer un

¹ <http://vision.stanford.edu/Science-2015-Turing.pdf>

² <https://www.openai.com/about/>

programme spécifique sur cette thématique¹. Cela implique des défis scientifiques mais aussi des problèmes pour former les générations futures à ces techniques, pour les démythifier et les comprendre. Un autre défi important pour ce domaine est souligné dans le rapport du département de la défense US². Il s'agit, notamment pour les technologies d'apprentissage profond (*deep learning*), de satisfaire aux propriétés souhaitables pour tout logiciel critique : preuve, maintenance, évolutivité, robustesse. Ces aspects posent des problèmes scientifiques complexes et nécessitent des collaborations entre des communautés scientifiques distinctes. L'importance de l'accès aux données et les enjeux économiques pour ces thématiques liées à l'apprentissage automatique fait que le secteur industriel économique et notamment les GAFA (Google, Apple, Facebook, Amazon) américains proposent une recherche de pointe dans ces domaines.

Robotique : la robotique est un secteur qui a également énormément évolué dans la dernière décennie, d'une part, en raison des progrès technologiques notamment sur les capteurs et, d'autre part, en raison des enjeux économiques et sociétaux qui ont suscité de nombreux investissements. Côté US, on peut citer les travaux du MIT sur le robot guépard³ mais aussi le robot humanoïde proposé par Google / Alphabet et Boston Dynamics⁴ capable de courir dans les bois. Concernant les perspectives, nul doute que les nanorobots seront au cœur de nombreux progrès scientifiques et technologiques dans les prochaines années.

2.4 Les États-Unis, un acteur majeur de la recherche en IA

Les États-Unis sont ainsi, comme d'ailleurs dans la plupart des domaines scientifiques, un acteur majeur de la recherche en intelligence artificielle. Le principal atout du pays est sans doute sa capacité à :

- attirer les meilleurs scientifiques et les talents les plus prometteurs,
- proposer de grands programmes et laboratoires thématiques en liaison avec ses agences (NSF⁵, NASA⁶, DARPA⁷, NIH, ...) et
- promouvoir l'innovation à travers un écosystème très efficace impliquant le monde académique et le monde industriel.

Il est également important de souligner la perméabilité des domaines scientifiques qui fait que l'innovation scientifique et technologique se retrouve souvent à la convergence de plusieurs domaines. L'IA a ainsi

¹ <http://www.darpa.mil/program/explainable-artificial-intelligence>

² Randolph Franklin

³ https://www.youtube.com/watch?v=_luhn7TLfWU

⁴ <https://www.theguardian.com/technology/video/2015/aug/18/humanoid-robot-atlas-boston-dynamics-run-video>

⁵ <https://www.nsf.gov/funding/programs.jsp?org=IIS>

⁶ <https://ti.arc.nasa.gov/tech/dash/physics/quail/>

⁷ <http://www.darpa.mil/tag-list?tt=73&type=Programs>

bénéficié des progrès récents dans le domaine du *big data* et du calcul haute performance (HPC).

Nul doute que la convergence des mondes numérique et physique avec le développement de l'Internet des objets et des véhicules autonomes sera la source de nombreux progrès où l'intelligence artificielle jouera un rôle majeur.

D'autres domaines comme la cybersécurité sont au cœur de la politique scientifique dans le numérique aux États-Unis avec des connexions fortes dans le monde de l'intelligence artificielle pour détecter et contrer les attaques.

Dans tous ces domaines les collaborations scientifiques entre la France et les États-Unis sont une réalité. Inria possède par exemple quarante équipes associées entre des équipes françaises et américaines. Le CNRS possède un Groupe de recherche international (GDRI) impliquant les États-Unis dans le domaine de l'intelligence artificielle et un laboratoire commun (UMI) avec Georgia Tech.

3. La course industrielle à l'armement

Naturellement, les promesses de croissance économique ouvertes par cette nouvelle rupture technologique ne pouvaient laisser indifférent l'écosystème des grands acteurs privés de l'innovation aux États-Unis.

3.1 Les reconfigurations internes des géants industriels

Il est notable que 2016 aura été l'année durant laquelle les géants industriels technologiques, notamment de la côte ouest, ont opéré une série de réorganisations internes destinées à mettre en avant des services de *cloud computing* spécialement conçus pour les avancées en intelligence artificielle.

C'est ainsi que par exemple Google/Alphabet vient de recruter Fei-Fei Li, directrice des *Artificial Intelligence and Vision Labs* de Stanford pour diriger son groupe d'apprentissage automatique dans le *cloud*, lequel inclura rapidement des services en reconnaissance d'image, reconnaissance vocale, traduction automatique ou compréhension du langage naturel. Microsoft, Amazon ou IBM ont tous engagé des restructurations similaires.

Une des raisons pour focaliser ce type d'effort réside dans la prédiction selon laquelle, le *cloud* étant promis à constituer une part dominante de la génération de leur chiffre d'affaires, l'infusion de cette partie de leur activité par les technologies d'intelligence artificielle et d'apprentissage automatique constitue un avantage comparatif décisif.

Plus encore, il s'agit pour ces sociétés de faire percoler ces technologies dans l'ensemble des divisions et produits associés des susdites sociétés. Microsoft vient par exemple de créer une division entière au sein de

laquelle son vice-président exécutif pour la recherche et la technologie, Harry Shum, conduit plus de cinq mille ingénieurs et informaticiens à travailler à la diffusion des technologies d'intelligence artificielle dans l'ensemble des produits de la société (le moteur de recherche Bing, l'assistant personnel Cortana ou les activités robotiques de la société en constituent les exemples les plus notables). Comme souvent, Google/Alphabet avait été un précurseur en la matière avec sa division Google Brain (portée par Andrew Ng, désormais à Baidu) qui travaille avec l'ensemble des autres divisions de Google/Alphabet. De la même manière, un ingénieur de Facebook sur cinq travaille désormais à partir des résultats en apprentissage automatique du pourtant récent *Applied Machine Learning Group*.

Il n'est d'ailleurs pas inintéressant de noter que la rareté relative des talents en intelligence artificielle (à l'inverse de ceux en programmation classique) conduit les grands industriels technologiques de la Silicon Valley à injecter des ressources significatives dans la formation (et la rétention) de ses employés : Google/Alphabet a monté un programme de formation interne en apprentissage automatique, et Facebook offre des formations similaires tout en construisant des parcours internes de transfert vers des missions à plein temps de recherche dans le domaine exclusif de l'intelligence artificielle.

3.2 Les stratégies d'acquisition

Une autre voie d'immersion pour les entreprises industrielles réside naturellement dans le déploiement d'une stratégie d'acquisition de sociétés plus petites, d'autant qu'elle s'accompagne souvent d'une volonté de rétention des personnels clés de ces dernières (ce qui se désigne par le terme d'*acqui-hire*).

Les chiffres parlent d'eux-mêmes : on passe d'un volume global d'acquisition de start-up spécialisées en intelligence artificielle de sept en 2012 pour tout le territoire américain, à trente-trois en 2014, puis trente-sept en 2015 et quarante-deux simplement pour les trois premiers trimestres de 2016.

Là encore, Google/Alphabet est à la fois un précurseur en la matière, puisqu'il a acquis dès 2013 la start-up d'apprentissage automatique issue de l'Université de Toronto DNNresearch, puis surtout en 2014 pour 600 M\$ la société britannique DeepMind Technologies, connue pour avoir enfanté un programme capable de battre le meilleur joueur mondial de go, et celui qui pratique cette stratégie avec le plus d'intensité et de constance (onze acquisitions au total, dont les plus récentes sont la start-up de recherche visuelle Moodstock et la plateforme de langage naturel Api.ai).

Mais ses concurrents directs ne sont pas en reste : Intel a acquis cinq start-up (dont trois en 2016), de même qu'Apple (deux en 2016), Twitter quatre (dont la très prometteuse start-up britannique de traitement d'image

Magic Pony pour 150M\$) et Salesforce, qui a démarré son effort en 2016, trois.

De manière intéressante, on voit désormais apparaître dans la course des entreprises industrielles non initialement liées dans leur cœur technologique à l'informatique et à l'Internet, le meilleur exemple en étant General Electric (deux acquisitions en novembre 2016)

3.3 A l'autre bout du spectre, l'expansion de l'écosystème de start-up en intelligence artificielle

Dans un marché de l'investissement en capital-risque sinon atone, du moins plafonnant après les records de l'année précédente, les start-up liées à l'intelligence artificielle et à l'apprentissage automatique connaissent *a contrario* une envolée spectaculaire de leur attractivité financière. Le troisième trimestre 2016 est à cet égard archétypal : une croissance impressionnante tant en volume global (705 M\$) qu'en nombre d'accords (71), soit une hausse respective de 16 % et 22 % par rapport au trimestre précédent, avec une tendance qui continue à accélérer.

L'élément le plus frappant réside cependant sans doute dans le degré inégalé de transversalité sur des verticales sectorielles par les start-up liées à l'intelligence artificielle, lesquelles attaquent désormais tous les marchés : le langage, la vision, l'automobile, la cybersécurité, la robotique, le marketing, l'intelligence économique, la génération de texte, l'Internet des objets, le commerce, les technologies financières, l'assurance, l'agriculture ou l'imagerie satellitaire. Mais la diversité de champs applicatifs au sein de ces marchés est tout aussi spectaculaire : si on considère le secteur qui a généré ces deux dernières années le plus d'investissement en capital-risque pour des start-up en intelligence artificielle, celui de la santé, on trouve aussi bien des sociétés qui travaillent sur la gestion du risque que sur la recherche génétique, la nutrition, l'imagerie médicale, le diagnostic, le monitoring individuel, la gestion des urgences hospitalières, la découverte de médicaments, l'oncologie ou la santé mentale. L'intelligence artificielle est donc bien devenue un marché total.

2. En Chine

Note réalisée avec le concours du service scientifique de l'Ambassade de France en Chine

L'**intelligence artificielle** (IA) est une discipline scientifique recherchant des méthodes de résolution de problèmes à forte complexité logique ou algorithmique. Par extension elle désigne, dans le langage courant, les dispositifs imitant ou remplaçant l'humain dans certaines mises en œuvre de ses fonctions cognitives.

Historiquement, elle trouve son point de départ avec les travaux de Turing dans les années 1950 et a suivi l'évolution de l'informatique. Le développement des technologies informatiques et en particulier l'explosion récente des puissances de calculs et de la capacité de capter et manipuler des quantités importantes de données ont permis, à l'aide d'algorithmes théorisés dans les années 1970, la réalisation de programmes informatiques surpassant l'homme dans certaines de ses capacités cognitives emblématiques.

Le sujet se trouve à l'interface de nombreuses disciplines. L'informatique est au centre mais en lien avec les modèles mathématiques sous-jacents - réseaux et statistiques pour l'analyse de données massives -, la biologie et en particulier les neurosciences, au moins comme source d'inspiration pour affiner les architectures, les sciences sociales comme les sciences cognitives ou la linguistique, mais aussi souvent les sciences liées aux applications concernées, de la ville intelligente à la médecine en passant par l'éducation ou l'automobile.

La Chine se trouve au cœur de l'évolution récente de ces techniques et déploie d'importants moyens pour devenir leader dans le domaine.

Parmi ses atouts, mentionnons l'importance de la part du PIB consacrée à la recherche et la mise en avant de la recherche autour de l'IA ; sa position dominante dans le classement du TOP 500 (deux supercalculateurs en tête de cette liste) ; la collusion État/géants du net et de l'informatique/instituts de recherche/universités/start-up ; un marché très important et friand des avancées potentielles du secteur ; une opinion publique peu préoccupée par les questions philosophiques soulevées ni par les questions concernant la protection des données qui sont des enjeux importants pour l'évolution et les applications de l'IA.

Ces avantages et les succès chinois annoncés en la matière sont à nuancer par certains handicaps : politique de communication parfois un peu arrogante et certaines fois déconnectée de la réalité (que ce soit pour la hauteur effective des financements ou la qualité des résultats obtenus) ; choix parfois guidés par une volonté plus de coller aux « indicateurs » que de faire progresser la science ; la plupart des progrès visibles reposent encore sur des architectures conçues par des scientifiques occidentaux ; existence de

multiples réseaux parfois concurrents se livrant à une compétition pas toujours très constructive (cela oblige d'ailleurs pour se faire une idée à identifier les différents réseaux) ; la pluridisciplinarité qui semble requise pour des avancées sérieuses en IA n'est pas dans la tradition scientifique chinoise ; et puis, mais ce n'est pas propre à la Chine, il y a une nette tendance à qualifier d'IA tout algorithme un peu original et pas trop bête.

Généralités

Le 13^e plan quinquennal comprend une liste de quinze « nouveaux grands projets - innovation 2030 » qui structurent les priorités scientifiques du pays et correspondent chacun à des investissements de plusieurs milliards d'euros. Parmi ces quinze projets, on trouve un projet de « Recherche sur le cerveau » et des projets d'ingénierie intitulés « Mega données », « Réseaux intelligents » et « Fabrication intelligente et robotique ». Si le terme « intelligence artificielle » n'apparaît pas explicitement, il est clairement sous-jacent dans ces quatre projets, ce qui signifie qu'un effort considérable va être développé dans cette direction.

À la manière locale, certains chercheurs présentent pour la communauté scientifique chinoise une feuille de route ambitieuse :

1. Actuellement, développement de machines ayant une intelligence par domaine ("*Artificial Narrow Intelligence*" - ANI) (reconnaissance d'objets, compréhension du langage naturel, assistant personnel...). Ces techniques pouvant être appliquées dans un grand nombre de domaines, l'objectif est que la Chine soit leader mondial (en 2020 ?).

2. Vers 2025-2030, développement de systèmes ayant une intelligence générale, comparable à celle du cerveau humain (« *Artificial General Intelligence* » - AGI). Ici aussi, l'ambition de la Chine est d'être la première à disposer d'un tel système.

3. Puis le développement d'intelligences suprahumaines. C'est d'après ces chercheurs la seule solution pour résoudre les problèmes de la planète - les intelligences humaines, trop limitées, ne peuvent appréhender ce type de problème dans leur ensemble, et donc sont incapables de proposer des solutions efficaces.

Les recherches envisagées ne sont pas seulement théoriques, les applications multiples sont un moteur important : détection des émotions, interface homme-machine, analyse d'images, contrôle de drones, de robots ou d'avatars, interface en langage naturel, automobiles autonomes, etc. Les systèmes permettront aussi de faire l'analyse des *big data*. L'exemple suivant ne semble pas gêner les collègues qui l'évoquent : « Piloté par le gouvernement et l'organisme central de planification, le dispositif de notation de la population devrait récupérer automatiquement les informations sur les citoyens d'ici à 2020. Il scrutera les activités en ligne, etc, pour générer un score individuel. Il semble que si un seuil est dépassé,

l'individu concerné se verra privé d'un certain nombre de droits et de services ».

Une expérimentation a lieu à Suining¹. D'autres dispositifs d'évaluation sont étudiés, comme le Sesame Credit du distributeur en ligne Alibaba (<http://www.bbc.com/news/world-asia-china-34592186>).

Les réseaux du Service pour la Science et la technologie de ce Poste diplomatique ont pu identifier quelques unités saillantes dans le paysage mouvant travaillant autour de ce sujet, en se reposant essentiellement sur les retours de quatre missions (une mission Découverte Chine sur l'informatique neuro-mimétique, une mission d'experts sur l'IoT, une mission d'experts sur le calcul haute performance et une mission d'experts sur le *big data*).

Cerveau

Ainsi la recherche en intelligence artificielle est en partie portée par le projet « cerveau » (qui comporte bien sûr par ailleurs une importante composante plus strictement étiquetée « sciences de la vie »), à travers le pôle « *brain-like/ brain-inspired computing* » qui est en train de se constituer. Il implique en particulier trois acteurs majeurs de la recherche : la Chinese Academy of Sciences (CAS), l'Université Tsinghua et l'Université de Pékin (Beida) qui pour l'instant ont développé leurs propres réseaux. Le programme « *CASIA brain* » est une collaboration entre six instituts de l'Académie des sciences chinoise qui implique, entre autres, l'institut d'automatique (CASIA) et l'institut de technologie informatique (ICT-CAS). De son côté, l'Université Tsinghua pilote le *Center for Brain-Inspired Computing Research* (CBICR) qui relie des partenaires dans toute la Chine et se présente comme résolument « transdisciplinaire » au moins au sein de l'informatique. Le groupe de l'Université de Pékin semble moins structuré nationalement mais entretient d'importantes coopérations avec les pays anglo-saxons. Dans tous ces cas, l'idée directrice est de travailler simultanément sur des aspects *hardware* et logiciel pour concevoir des circuits susceptibles de fonctionner « comme le cerveau » pour apprendre à résoudre les problèmes posés.

Dans la veine de l'informatique neuro-mimétique, on relève aussi des travaux intéressants à Hangzhou (Projet Darwin, *College of Computer Science*, Université de Zhejiang) et à Shanghai (*Center for Brain Like Computing and Machine Intelligence* de l'Université Shanghai Jiaotong).

¹ Cf. https://www.washingtonpost.com/world/asia_pacific/chinas-plan-to-organize-its-whole-society-around-big-data-a-rating-for-everyone/2016/10/20/1cd0dd9c-9516-11e6-ae9d-0030ac1899cd_story.html?utm_term=.71ca38d649e1.

Big data

Parfois au sein des mêmes institutions, les laboratoires d'informatique (leur département « logiciel »), parfois au sein de pôles constitués autour des enjeux dits « *big data* » ou « IoT » développent des outils et intègrent des méthodes avec des objectifs de même nature – du moins pour les objectifs à court terme.

Outre les laboratoires des grandes institutions pékinoises (*CAS Key Lab of Network Data Science and Technology* ou *Institute for Data Science Tsinghua*), on relève le laboratoire *Pattern Recognition & Intelligent System Lab.* et le laboratoire du logiciel de télécommunication intelligent et multimédia de la *Beijing University of Post and Telecommunications (BUPT)*, la *School of Computer Software* à Tianjin, le *National Key Laboratory for Novel Software Technology* de l'Université de Nankin, le *Shenzhen Institute of Advanced Technology* de la CAS, ou encore le *Shanghai Advanced Research Institute* de la CAS (centré sur l'IoT), l'*Institute of Media Computing* de l'Université Fudan, également à Shanghai.

Géants privés

Les enjeux sont nationaux et les financements suivent. Mais les entreprises privées sont aussi des moteurs puissants du secteur.

En tête, l'entreprise *Baidu* (qui a développé le moteur de recherche chinois, site le plus consulté en Chine et cinquième au niveau mondial, indexant près d'un milliard de pages, cent millions d'images et dix millions de fichiers multimédia) communique beaucoup sur le sujet, consacre une part conséquente de sa recherche au sujet (*Institute of deep learning, Big data lab,...*) et, comme les géants américains, dispose d'un flux de données permettant d'envisager des applications dans de nombreux domaines.

L'entreprise considère, comme ses concurrents, que l'IA est son prochain challenge comme solution clé pour des applications en vision, parole, traitement du langage naturel et sa compréhension, génération de prédictions et de recommandations, publicité ciblée, planification et prise de décision en robotique, conduite autonome, pilotage de drones,... Elle travaille en étroite relation avec de nombreuses universités et start-up.

Les résultats algorithmiques de l'*Institute of Deep Learning* de Baidu sont impressionnants et du meilleur niveau mondial, malgré son existence récente. Mentionnons pour donner une idée, la capacité de leurs algorithmes à retrouver une image dans une base de données de 10 milliards d'images en moins d'une seconde. Ils ont aussi de bonnes performances sur les *benchmarks ICDAR* (1^{er} sur 5/8 évaluations parmi 4 tâches).

Sur *FDDB (Face Detection Data Set and Benchmark)*, voir <http://www.cs.umass.edu/fddb/>) et sur la base de données de visages *LFW*, ils progressent vite (8 % d'erreurs en décembre 2015, 2,3 % en septembre 2016 et bientôt 1 %). L'entreprise annonce aussi la meilleure précision sur la

collection de benchmarks KITTI (<http://www.cvlibs.net/datasets/kitti/>) orientés pour la conduite de voitures autonomes. Baidu développe aussi des applications de reconnaissance d'image pour la plateforme de services *Baidu Nuomi* : une application permet par exemple de reconnaître le restaurant (et le plat) en prenant une photo de nourriture dans un restaurant.

Il n'est pas évident d'évaluer la maîtrise théorique des ingénieurs mais l'entreprise montre une incontestable efficacité pour implémenter rapidement les dernières innovations du secteur.

Les autres géants chinois du net, comme Alibaba (distribution) ou Tencent (réseaux sociaux), tirent eux aussi dans la même direction : développement et diffusion grand public d'applications plus ou moins convaincantes mais manifestement exploitant des techniques d'IA un peu évoluées, même s'ils semblent moins présents dans la recherche « fondamentale ». Huawei (telecom, téléphones), entreprise qui accorde une grande importance à la recherche fondamentale et dont la recherche s'internationalise rapidement (pôle mathématique implanté en France il y a deux ans) a mis en avant début janvier 2017 un concept de téléphone intelligent, dont on ne peut encore savoir s'il ira effectivement plus loin que ceux de ses concurrents.

Plus globalement concernant les données, les pôles universitaires peuvent aussi compter sur le support des industriels comme, par exemple, National Grid, China Mobile, China Unicom, China Mobile, Shanghai Meteorological Bureau ou Environmental Monitoring Center.

Innovation et start-up

Le paysage chinois de l'innovation - incitation à très grande échelle pour soutenir l'innovation privée, en particulier par la multiplication de petites structures - conjugué à l'effet de mode « attractif » du secteur IA fait éclore dans tout le pays et en particulier aux abords des (voire dans les) universités, des entreprises cherchant à exploiter ce filon, souvent en proposant des services originaux ; sous l'œil (attentif et bienveillant) de l'État et des géants. Ces initiatives ne sont pas toutes scientifiques. Il s'agit souvent d'abord de marketing, mais pas toujours ou pas seulement.

On peut mentionner par exemple la société Cambricon, très proche de l'ICT de la CAS : le professeur Chen Yunji continue de développer des accélérateurs neuronaux (à l'ICT de la CAS) de la famille DianNao à partir de travaux réalisés en coopération avec des chercheurs du CEA (Marc Duranton) et d'Inria (Olivier Temam) ; son frère Chen Yianshi est maintenant le P-DG d'une start-up « Cambricon Technologies Corporation Limited » fondée en mars 2016, comptant 100 employés en septembre 2016 et ayant un objectif de 200 pour fin 2016, qui travaille sur des accélérateurs neuronaux inspirés fortement de la famille DianNao. Le projet de la société est d'utiliser ces accélérateurs pour permettre le traitement intelligent des images et des

signaux embarqués (au lieu d'architectures GPU / DSP actuelles) améliorant les performances énergétiques d'un facteur 4000. Le chercheur Marc Duranton qui a eu accès (limité) à la société estime que cette perspective est prometteuse.

Sur le marché des robots de service, intégrant de l'IA en particulier pour la partie reconnaissance vocale, mentionnons par exemple Yunji Technology (robots déployés dans des hôtels à Pékin et Shanghai), iFLYTEK (moteur de reconnaissance vocale) ou Ninebot Inc (Transport intelligent).

Certaines sociétés françaises du secteur, comme par exemple *Air Visual* (suivi et prédiction de la pollution de l'air) ou *WosomTech* (reconstruction 3D prédictive ; approchée par *Lenovo* récemment) ont une expérience en Chine et connaissent un peu le paysage de l'intérieur.

Personnes ressources à privilégier

Mentionnons tout d'abord les trois principaux acteurs de ce qui devrait constituer le pôle « *brain-like/inspired computing* » du projet de recherche sur le cerveau Il s'agit de HUANG Tiejun, SHI Luping et XU Bo, travaillant respectivement à l'Université de Pékin, à l'Université Tsinghua et à l'Institut d'Automatique de la CAS.

Le NLPR (*National Laboratory of Pattern Recognition*) dirigé par TAO Jinhua est l'un des laboratoires de l'Institut d'Automatique de la CAS. Il est notable qu'il héberge le principal laboratoire franco-chinois du domaine, le LIAMA (Laboratoire d'Informatique, d'Automatique et de Mathématiques Appliquées) du domaine. L'intelligence artificielle, et en particulier ce qui tourne autour du « *brain-inspired computing* » est l'un des sujets mis en exergue lors de la dernière rencontre organisée par le LIAMA (Pékin, Octobre 2016).

Le Professeur TAO Jinhua a encadré au NPLR la thèse d'un étudiant qui a obtenu récemment un « *one-million prize* » décerné par l'entreprise *Baidu* - pour son travail sur la reconnaissance vocale et la production de parole. L'une de ses étudiantes actuelles travaille sur la reconnaissance des émotions dans la voix humaine. C'est nous semble-t-il un contact précieux dans ce secteur.

ZHOU Zhihua est directeur adjoint de l'équipe LAMDA (Learning And Mining from DATA), affiliée au National Key Laboratory for Novel Software Technology et au Department of Computer Science & Technology. C'est un chercheur qui a aujourd'hui une véritable envergure internationale et qui est une valeur montante de l'apprentissage statistique en Chine.

Les Hub French Tech récemment labellisés à Pékin, Shanghai et Shenzhen peuvent constituer d'excellents points d'entrée pour accéder aux start-up françaises du secteur ayant des intérêts en Chine.

QIU Xipeng : vice dean du Big Data School de l'Université de Fudan.

3. Au Japon

Note réalisée avec le concours du service scientifique de l'Ambassade de France au Japon

L'intelligence artificielle est considérée comme l'élément clé de la révolution numérique au Japon. Le gouvernement japonais en a fait le socle de sa nouvelle stratégie en science, technologie et innovation, qui vise à mettre en place une « société 5.0 », société superintelligente et fer de lance à l'échelle mondiale. Il a annoncé une vague d'investissements massifs dans le domaine, avec 27 milliards de yens pour la seule année 2016, à travers la création de centres de recherche et technologies dédiés à ce domaine. Les partenariats publics-privés entre ces nouveaux centres et les grands groupes japonais se multiplient, afin d'exploiter le potentiel de création de valeur que constitue l'intelligence artificielle sur des applications ciblées.

1. L'intelligence artificielle, au cœur de la stratégie japonaise

1.1. L'intelligence artificielle (AI) dans le « 5e Plan-cadre de la science et de la technologie » et dans le « Livre blanc 2016 sur la science et la technologie »

Les plans-cadres pour la science et de la technologie, élaborés tous les cinq ans depuis 1996 par le Conseil pour la Science, Technologie et Innovation (CSTI, dépendant directement du Cabinet Office, i.e. Les services du Premier Ministre japonais), définissent les priorités de l'État dans les domaines de la science, de la technologie et de l'innovation.

Entré en vigueur le 1^{er} avril 2016 (le début de l'année fiscale 2016), le 5^e Plan-cadre¹ repose sur un nouveau concept de « Société 5.0 », société ultra-intelligente, qui tirerait le meilleur bénéfice des opportunités offertes par les technologies numériques afin de redynamiser le secteur industriel, mais également dans l'optique de réformer la structure même de la société et assurer une prospérité inclusive, dans laquelle les citoyens bénéficieraient de services de haute qualité et adaptés à leurs besoins selon leur âge, sexe, région ou langue. Pour mettre en œuvre cette société fer-de-lance à l'échelle mondiale, le Japon compte tirer profit de l'expansion très rapide des technologies de l'information et de la communication, parfois considérée comme la « 4^e révolution industrielle », et s'appuyer sur trois piliers : l'Internet des objets (IoT), le *big data* et l'intelligence artificielle (IA).

L'intelligence artificielle ouvre la voie à des développements très attendus dans le domaine des transports, de la communication, de la traduction automatique ou de la robotique, notamment à l'horizon des Jeux Olympiques et paralympiques de 2020, que le Japon envisage comme une vitrine technologique pour se présenter comme le pays leader en termes d'innovation.

¹ http://www8.cao.go.jp/cstp/english/basic/5thbasicplan_outline.pdf

Après la publication du 5^e Plan-cadre pour la Science et la Technologie par le CSTI, le Ministère japonais de l'Éducation, de la Culture, des Sports, des Sciences et de la Technologie (MEXT) a rédigé son livre blanc 2016 pour la science et la technologie, intitulé « *Vers la société ultra-intelligente mise en œuvre par l'IoT, le big data et l'IA – pour que le Japon soit un précurseur mondial* », adopté en Conseil des ministres le 20 mai 2016. Ce rapport signale l'insuffisance au Japon de la formation et de la recherche fondamentale en informatique, qui devraient venir soutenir le développement des technologies de l'information et de la communication.

Enfin, à l'occasion de l'organisation du G7 au Japon, le Japon a pris l'initiative d'organiser une réunion ministérielle G7 consacrée aux Sciences et technologies de l'information et de la communication (format qui n'avait pas été mis en œuvre depuis 20 ans). Pour faire suite à la tenue de cet événement en avril à Takamatsu, les ministres en charge des Sciences et Technologies de l'Information (ICT) se sont accordés dans leur déclaration conjointe à promouvoir la R&D pour les technologies ICT, notamment dans le domaine de l'intelligence artificielle. Ils ont mis l'accent sur l'importance de mettre en place des politiques adaptées, permettant de prendre en compte l'impact sociétal et économique colossal de ces technologies.

1.2. Mise en place de deux comités pour le suivi de ces stratégies

Création d'un Comité de délibération sur l'IA et la société humaine

Le 12 avril 2016, à l'issue de la réunion régulière du Conseil des ministres, Mme Aiko Shimajiri, Ministre chargée de la politique de la science et de la technologie, a annoncé la mise en place d'un « **Comité de délibération sur l'AI et la société humaine** » au sein du Cabinet Office. Il s'agit de la première entité gouvernementale ayant pour mission d'étudier les enjeux liés à l'IA selon cinq points de vue, à savoir : l'aspect éthique, l'aspect légal, l'aspect économique, l'aspect social et l'aspect R&D. Ce Comité, présidé par Mme Yuko Harayama, membre exécutif permanent du CSTI du Cabinet Office, est composé de 11 experts. La première réunion s'est tenue le 30 mai 2016 en présence de la ministre Shimajiri. Le Comité s'est réuni sur une base mensuelle jusqu'en septembre 2016 pour analyser les activités nationales et internationales et approfondir la problématique. Il a choisi de se baser sur des cas d'application précis, mettant en œuvre des technologies qui devraient voir le jour à court terme : le véhicule autonome, l'automatisation de l'appareil de production et la communication homme/machine. Il souhaite également engager un débat avec le grand public (par le biais essentiellement de séminaires ouverts et de questionnaires en ligne).

Le Comité devra remettre avant la fin de l'année 2016 ses conclusions, qui seront prises en compte dans la nouvelle Stratégie globale pour la Science, Technologie et Innovation qui sera publiée en juin 2017. Les discussions au niveau international en la matière sont prévues à partir de

2017. Mme Harayama a présenté à Paris le 17 novembre 2017, dans le cadre du *Technology Foresight Forum 2016* de l'OCDE¹ dédié à l'intelligence artificielle, les premières réflexions du comité :

- Éthique : le citoyen peut-il accepter d'être manipulé pour modifier ses sentiments, convictions ou comportements, et d'être catégorisé/évalué, sans en être informé ? Quel impact aura le développement de l'IA sur notre sens de l'éthique et les relations entre les hommes et les machines ? Dans la mesure où elle étend notre temps, notre espace et nos sens, l'IA viendra-t-elle affecter notre conception de l'humanité, notamment notre conception des facultés et des émotions humaines ? Comment évaluer les actions et la création à partir de l'IA ?

- Légal : comment trouver le juste équilibre entre les bénéfices du traitement des *big data* par l'IA et la protection des informations personnelles ? Est-ce que les cadres légaux existants pourront s'appliquer aux nouvelles problématiques soulevées par l'IA ? Comment clarifier la responsabilité dans le cas d'incidents impliquant de l'intelligence artificielle (par exemple pour le véhicule autonome) ? Quels sont les risques d'utiliser l'IA ? De ne pas utiliser l'IA ?

- Économique : comment l'IA va-t-elle changer notre manière de travailler ? Quelle politique nationale mettre en place pour favoriser l'utilisation de l'IA ? Comment l'IA va-t-elle modifier le monde de l'emploi ?

- Sociétal : comment réduire les divisions liées à l'IA et répartir de manière équitable le coût social de l'IA ? Y-a-t-il une pathologie ou des conflits de société que peut potentiellement engendrer l'IA ? Peut-on assurer la liberté d'utiliser ou non l'IA et assurer le droit à l'oubli ?

- Éducation : quelle politique nationale mettre en place pour faire face aux inégalités que ne manqueront pas de causer l'utilisation l'IA dans le domaine de l'éducation ? Comment développer notre capacité à exploiter l'IA ?

- R&D : quelle R&D développer pour l'IA en respect de l'éthique, de la sécurité, de la protection de la vie privée, de la transparence, de la contrôlabilité, de la visibilité, de la responsabilité ? Comment rendre disponible l'information liée à l'IA de manière à ce qu'un utilisateur puisse prendre la décision d'utiliser ou non l'IA ?

Le comité cherche notamment à approfondir trois voies pour définir des politiques adaptées : la coévolution de la société et de la technologie ; la recherche d'un équilibre entre les bénéfices (services personnalisés à coût abordable) et les risques liés à l'IA (discrimination, perte de protection des données à caractère privé, perte d'anonymat) ; la définition des limites de la prise de décision automatisée.

¹ <http://www.oecd.org/sti/ieconomy/Yuko%20Harayama%20-%20A1%20Foresight%20Forum%20Nov%202016.pdf>

Un Conseil portant la stratégie des technologies liées à l'IA

Par ailleurs, dans le cadre du « Dialogue privé-public pour les investissements du futur », le Premier ministre Abe a présenté son projet de définir d'ici à la fin de l'année la feuille de route présentant les objectifs de la recherche sur l'AI et ses applications industrielles et de mettre en place un « **Conseil de la stratégie des technologies liées à l'AI** ».

Ce Conseil, dont la réunion inaugurale a eu lieu le 18 avril 2016, est présidé par M. Yuichiro Anzai, Président de la JSPS (agence de financement de la recherche du MEXT consacrée à la recherche fondamentale), accompagné d'un conseiller, M. Kazuo Kuyma, membre exécutif permanent du CSTI, et composé de deux représentants du Keidanren (syndicat patronal des entreprises japonaises), des présidents de deux universités (Université de Tokyo et Université d'Osaka) et des présidents de cinq grands instituts de recherche et agences de financement : le NICT (*National Institute for Information and Communication Technologies*, dépendant du Ministère des affaires intérieures et de la communication, MIC), le RIKEN (principal centre de recherche pluridisciplinaire du MEXT), l'AIIST (*Advanced Institute for Science and Technologies*, centre de recherche multidisciplinaire du Ministère de l'Économie et de l'Industrie, METI), la JST (l'une des deux agences de financement du MEXT, orientée vers les projets de recherche appliquée) et la NEDO (agence de financement dépendant du METI).

Ce Conseil servira de quartier général de la R&D des technologies de l'AI et de leurs applications industrielles en regroupant les trois ministères impliqués dans l'AI : le MIC, le MEXT et le METI.

Deux Comités ont été mis en place sous ce Conseil : le « Comité de collaboration de recherche » (conseil des présidents des instituts de recherche et des agences de financement) et le « Comité de collaboration industrielle » qui se réunissent chacun mensuellement.

2. Mise en œuvre : les trois ministères impliqués mettent en œuvre leurs propres initiatives

Dans ce contexte, les trois ministères, respectivement en charge de la recherche, de l'industrie et des communications (MEXT, METI et MIC), se sont mobilisés avec une vitesse étonnante pour développer des initiatives dans ce domaine. Ils ont chacun annoncé la création d'un centre de recherche sur l'intelligence artificielle en 2016.

- Le **Ministère japonais de l'Éducation, de la Culture, des Sports, des Sciences et de la Technologie** (MEXT) a officiellement lancé son centre, le « AIP Center » (*Advanced Integrated Intelligence Platform Project Center*), le 1^{er} avril 2016, hébergé par le RIKEN (principal centre de recherche pluridisciplinaire du MEXT) et dirigé par le Professeur Masashi Sugiyama. Doté d'un budget de 1,45 milliards de yens pour l'année fiscale 2016, ce centre vise à construire une approche intégrée des technologies d'intelligence

artificielle/ *big data* /IoT et cybersécurité. L'acceptation du terme intelligence artificielle est prise au sens large, incluant la composante informatique (algorithme, réseaux, architecture matérielle), mais également mathématique (algèbre linéaire, probabilités et statistiques, optimisation), ingénierie (intégration dans des systèmes) et applications dans différents domaines scientifiques (biologie, physique, science des matériaux..). Les activités s'articulent autour de cinq objectifs :

- développer des technologies pour l'intelligence artificielle « fondamentale » (basé essentiellement sur le *machine learning* et le *deep learning*, avec un effort portant sur un apprentissage robuste et en temps réel sur des données pouvant être bruitées, hétérogènes, incomplètes...)

- contribuer à l'accélération de la recherche scientifique (analyse et synthèse automatique d'articles scientifiques, brevets, résultats d'expériences, design efficace de molécules et matériaux...)

- contribuer à des applications concrètes à fort impact sociétal : problématique des soins dans le contexte du vieillissement de la population, gestion des infrastructures (inspection des fissures dans les ouvrages de génie civil tel que les ponts, etc...) résilience aux catastrophes naturelles ;

- prise en compte des aspects éthiques, légaux et sociaux ;

- développement des ressources humaines.

Ce centre (pour lequel un nouveau bâtiment est en cours de construction, au centre de Tokyo), dirigé par le professeur Masahi Sugiyama de l'Université de Tokyo, est composé de 28 personnes, réparties entre deux groupes de recherche, l'un consacré aux technologies génériques, et l'autre aux technologies dédiées à certaines applications. Il cherche à accroître ses effectifs, notamment en accueillant des chercheurs étrangers pour des durées longues. Par ailleurs, une vingtaine d'entreprises, dont Toyota Motors, NEC, Sony Computer Science Laboratories, NTT ou start-up spécialisée dans le *Deep Learning Preferred Networks* ont annoncé la mobilité de plusieurs de leurs chercheurs dans ce centre. Le centre devrait rapidement monter en puissance, avec un budget de 10 millions de yens qui sera octroyé par le MEXT pour l'année 2017, et plusieurs dizaines de millions de yens attendus de la part des partenaires privés.

- **Le MEXT a également mis en place un laboratoire virtuel** réunissant les projets de recherche financés par la JST (l'une des deux agences de financement du MEXT, orientée vers les projets de recherche appliquée) sur l'intelligence artificielle, dit « **Network Labo** », dirigé par le Professeur Setsuo Arikawa, ancien président de l'Université de Kyushu.

Cinq projets de recherche sur l'AI sont déjà financés par la JST, pour une somme totale de 2,85 milliards de yens pour l'année fiscale 2016. La JST mettra en place trois nouveaux projets additionnels dans ce domaine, dotés d'un montant de 1,15 milliards de yens en 2016. Ces trois nouveaux projets relèvent respectivement de la médecine (identification de symptômes par

traitement de données massives d'imagerie médicale), de la conduite autonome (par traitement de mégadonnées issues de caméras et de radars à extrêmement haute fréquence installés dans les véhicules) et de la cybersécurité. Ces huit projets en intelligence artificielle seront regroupés et pilotés *via* le « Network Labo », ce qui permettra de mutualiser certaines des technologies de base développées dans le cadre de ces projets et de diversifier les applications des résultats de recherche.

Le « **Network Labo** » et l'**AIP Center** collaborent étroitement afin de mutualiser leurs ressources matérielles et humaines.

- **Le Ministère de l'Économie et de l'Industrie (METI)** a lancé également son centre dédié à l'IA le 1^{er} mai 2015 : l'**AIRC** (*Artificial Intelligence Research Center*), hébergé par l'**AIST** (*Advanced Institute for Science and Technologies*, centre de recherche multidisciplinaire du METI). Il regroupe 348 personnes en 2016, dont 269 chercheurs (parmi lesquels 95 chercheurs permanents, 16 détachés d'universités, 17 détachés d'entreprises et 41 professeurs étrangers invités) et repose sur 3 piliers : plateformes de données, compétences en informatique et « *business models* » clairs sur des applications ciblées (axées sur la production industrielle, notamment la robotique industrielle ; les grands secteurs à fort impact sociétal et économique tels que la santé, le tourisme ; les disciplines scientifiques utilisant de grandes masses de données).

Le centre est financé par METI et son agence de financement la NEDO pour un montant de 19,5 milliards de yens pour l'année 2016, ainsi que par des contrats directs avec des acteurs industriels sur des applications ciblées (l'AIRC a notamment mis en place un laboratoire commun avec l'entreprise NEC).

L'AIRC vise à intégrer l'approche orientée données (*machine learning*, *deep learning*, modèles statistiques) et l'approche orientée « connaissance », utilisant une ontologie du domaine d'application considéré. Il mène également des recherches sur le cerveau humain. Enfin, il porte un effort particulier sur les infrastructures de calcul (il va notamment faire l'acquisition d'un superordinateur d'une puissance de 130 PetaFlops : l'*AI Bridging Cloud Infrastructure* (ABCI) en 2017, pour un financement de 19.5 milliards de yens en provenance du METI).

- **Le Ministère des Affaires intérieures et Communication (MIC)** a créé un *cluster* de recherche dédié à l'intelligence artificielle, hébergé par son centre de recherche, le NICT (*National Institute for Information and Communication Technologies*). Ce *cluster* de recherche¹ se focalise sur l'IA inspirée par le cerveau humain, la reconnaissance vocale, la traduction automatique multi-langue ou l'analyse de la connaissance sociale. Il est doté d'un budget de 2,2 milliards de yens pour l'année 2016.

¹ Voir les pp. 6et 7 : http://www.nict.go.jp/en/data/nict-news/NICT_NEWS_1606_E.pdf

Si les trois ministères ont mis en place leurs centres de recherche en propre, ils se sont engagés à coordonner leurs actions, *via* le Conseil dédié à la stratégie des technologies liées à l'IA, à organiser des symposiums communs sur une base annuelle, à mettre en place un site web commun, et à mutualiser leurs ressources informatiques.

3. Les acteurs industriels investissent massivement dans le domaine

On estime que le marché de l'intelligence artificielle au Japon devrait passer de 3,7 milliards de yens en 2015 à 87 milliards en 2030, dont 30,5 milliards de yens dans le domaine du transport et 42 milliards en incluant la production industrielle pour le transport.

Plusieurs grands groupes japonais¹ ont lancé de grands programmes ou laboratoires dédiés à ce domaine, pour un montant d'investissement devant atteindre 300 milliards de yens sur les trois ans à venir.

De nombreux projets sont développés en perspective des Jeux Olympiques et Paralympiques de 2020, qui verront un afflux de visiteurs étrangers à Tokyo, motivant le développement de systèmes de traduction automatique, tels que ceux développés par Panasonic (qui se positionne également sur la robotique, la reconnaissance d'image et la reconnaissance spatiale, en collaboration avec le NICT et plusieurs universités japonaises, ainsi que dans ses centres de R&D situés à Singapour et dans la *Silicon Valley*) ou le *chatbot* multilingue de KNT-CT. Tokyo sera à cette occasion également la cible privilégiée d'attaques physiques ou cyber, ce qui incite Fujitsu Laboratories développer des systèmes intelligents de protection contre les attaques cyber, tandis que NEC met au point un système de reconnaissance faciale, NeoFace, basé sur l'IA.

Le secteur automobile draine également des investissements considérables, avec le projet « Robot Taxi » de DeNA et ZMP, ou le partenariat entre Honda et Softbank sur un projet d'IA pour l'assistance à la conduite automobile. Honda a par ailleurs annoncé en juin 2016 la création

¹ Les exemples donnés peuvent être approfondis sur les sites suivants :

<http://asia.nikkei.com/Business/Trends/Japan-tech-giants-double-down-on-AI> ;

<http://news.panasonic.com/global/stories/2016/45102.html> ;

<http://asia.nikkei.com/Tech-Science/Tech/Multilingual-chatbot-to-link-japan-hotels-foreign-tourists> ;

<http://asia.nikkei.com/Tech-Science/Tech/Cybersecurity-companies-adopting-AI-but-so-are-hackers> ;

<http://www.japantoday.com/category/technology/view/nec-launches-ai-software-that-searches-video-for-specific-individuals> ;

<https://robbotaxi.com/en/service/#servicevision> ;

<http://www.japantimes.co.jp/news/2016/06/02/business/corporate-business/honda-set-ai-research-base-tokyo/> ;

<http://asia.nikkei.com/Business/Trends/Japan-to-launch-initiative-for-AI-based-drug-discovery> ;

<http://www.japantoday.com/category/technology/view/ai-based-smart-energy-system-developed> ;

<http://www.roboticsbusinessreview.com/deep-learning-startup-pfn-partners-fanuc-save-japanese-manufacturing/> ;

<http://www.japantoday.com/category/technology/view/hitachi-develops-basic-ai-technology-that-enables-logical-dialogue-in-japanese> ;

<http://asia.nikkei.com/Business/Consumers/NEC-s-AI-system-knows-who-you-are-and-what-you-want>

d'un centre de R&D à Tokyo dédié à l'intelligence artificielle, Honda R&D Innovation Lab Tokyo, qui regroupera l'ensemble des activités de la société dans le domaine de l'IA. Toyota a de son côté investi 1 milliard de dollars sur l'IA dans son centre de la Silicon Valley aux États-Unis.

Dans le domaine de la santé, L'entreprise *Preferred Network* a lancé avec le Centre national du Cancer et l'AIRC un projet conjoint de recherche sur l'IA pour le diagnostic précoce et le traitement du cancer. Fujitsu et NEC ont décidé d'investir dans la découverte de nouveaux médicaments grâce à l'IA tandis que Sony Computer Science Laboratories développe avec l'AIP Center un système médical basé sur l'IA, permettant de recommander automatiquement des traitements adaptés à des symptômes donnés.

Dans le secteur de l'énergie, Hazawa Ando Corp un développe un système de gestion intelligent qui a pour but de réaliser un « Net Zero Energy Building » utilisant l'énergie solaire.

Enfin, dans le secteur de l'industrie de production, l'IA devrait redonner un second souffle aux usines japonaises. Canon, qui a annoncé le retour de sa production au Japon se base sur l'IA pour construire une usine entièrement automatisée à l'horizon 2018. FANUC a également investi massivement dans l'IA pour ses usines de production, *via* un investissement de 900 millions de yens en 2015 dans la start-up Preferred Networks.

De son côté, Hitachi s'est positionné sur le développement d'outils d'IA pour l'aide à la décision par les entreprises. NEC fournit également des solutions de suivi de clients pour le commerce de détail.

4. Développement de la collaboration scientifique franco-japonaise sur l'intelligence artificielle

Pour approfondir les opportunités de collaboration entre le Japon et la France dans ce domaine, le Service pour la science et la technologie de l'Ambassade de France au Japon a organisé, en partenariat avec l'Université de Tokyo (Policy Alternatives Research Institute), Elsevier et CEA Tech un événement de trois jours, les 11, 12 et 13 octobre 2016 sur l'apprentissage profond (*deep learning*) et l'intelligence artificielle.

Cet événement était composé d'un symposium et d'un programme de visites, dont celle des nouveaux centres d'intelligence artificielle du MEXT (mis en œuvre par le RIKEN) et du METI (mis en œuvre par l'AIST), celle du Policy Alternative Research Institute de l'Université de Tokyo (PARI) ainsi que du Laboratoire franco-japonais d'informatique¹. Lors de ces visites, les responsables de ces trois nouveaux centres ont montré une forte volonté de développer les collaborations internationales (échanges de chercheurs, partage de grandes bases de données..).

¹ Le JFLI, réunissant l'Université de Tokyo, l'Université Keio, le NII (National Institute of Informatics, dépendant du MEXT) côté japonais, le CNRS, Inria et l'Université Pierre-et-Marie-Curie côté français.

Le symposium a réuni le 12 octobre à l'Université de Tokyo près de 200 personnes, représentants du gouvernement et d'agences gouvernementales (CSTI, MIC, MEXT, JST), experts français et japonais, dans le domaine du *deep learning* et du *machine learning*, issus d'institutions de recherche publique ou d'entreprises.

Cet événement a mis en valeur la qualité de la collaboration existante entre la France et le Japon ainsi que l'existence d'un cadre favorable (programmes de financements, possibilités d'échanges *via* le JFLI), que les acteurs peuvent à mettre à profit pour développer plus encore cette coopération.

Les différentes présentations et discussions ont fait émerger plusieurs thèmes à fort potentiel pour cette coopération, tels que l'apprentissage par renforcement ou la question de l'accès et de l'utilisation.

Pour entretenir cet élan, un groupe de travail réunissant les acteurs clés du domaine en France et au Japon est en cours de création.

Des premières discussions ont également été lancées avec le CSTI, qui a manifesté son intérêt pour échanger avec les experts français sur les questions intelligence artificielle et éthique, en vue de préparer les recommandations pour le gouvernement japonais dans ce domaine.

Enfin, l'Ambassade de France au Japon organisera en 2017 un cycle de débats d'idées sur le sujet de l'intelligence artificielle, axés sur l'évolution des modes de travail, les *smart cities*, etc.

5. Conclusion

Le Japon mise fortement sur l'intelligence artificielle pour redynamiser son économie, notamment pour conserver, voire rapatrier, son appareil de production, *via* l'automatisation complète des usines.

Il s'agit de tirer profit du potentiel de l'intelligence artificielle (essentiellement sur les technologies du *deep learning*), de manière intégrée aux domaines de l'IoT et du *big data*.

Le fait que de nombreux acteurs académiques et industriels possèdent des infrastructures de calcul intensif (super-ordinateurs) est également un atout dans ce domaine.

Le gouvernement japonais accompagne cette transition de manière très étroite, avec des investissements publics massifs, mais également par la constitution de comités de réflexion sur l'impact de l'utilisation de l'intelligence artificielle sur la société et la nécessaire adaptation des réglementations.

4. Au Royaume-Uni

Note réalisée avec le concours du service scientifique de l’Ambassade de France au Royaume-Uni

Une recherche britannique en IA dynamique et soutenue par des infrastructures de mutualisation et de mise en réseau.

Plusieurs équipes recherche et centres d’innovation se sont développés au Royaume-Uni au cours des dernières années dans les domaines connexes de la robotique et de l’intelligence artificielle.

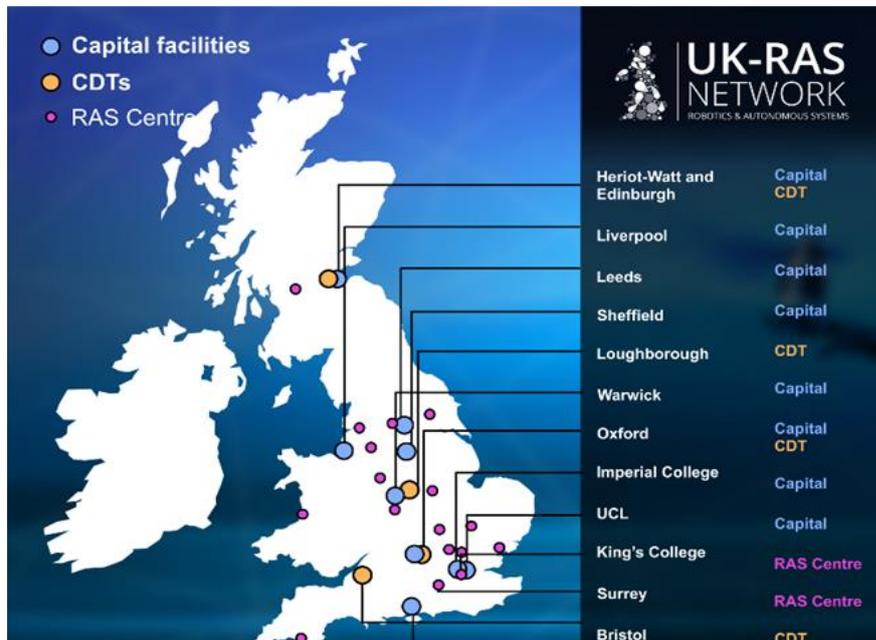
Afin de répondre aux besoins d’interdisciplinarité et de mutualisation de la recherche en intelligence artificielle, plusieurs organismes et réseaux se sont développés ces dernières années au Royaume-Uni.

C’est le cas du réseau britannique de robotique et systèmes autonomes de l’Engineering and Physical Sciences Research Council (EPSRC), l’EPSRC UK-RAS Network, créé en mars 2015.

L’EPSRC est une institution britannique qui fournit des financements pour la recherche et l’enseignement supérieur dans le domaine de l’ingénierie et des sciences physiques.

Le réseau EPSRC UK-RAS a pour mission d’assurer une collaboration entre acteurs académiques et industriels.

Il coordonne les activités de huit infrastructures (*Capital facilities*) et quatre centres de formation doctorale (*Centres for Doctoral Training – CDT*) sur la robotique et les systèmes autonomes, à travers le Royaume-Uni.



Crédits : EPSRC UK-RAS

Les universités participant au réseau EPSRC UK-RAS sont les suivantes :

Université d'Édimbourg et Université Heriot-Watt (**Edinburgh Centre for Robotics**), Imperial College London (**Hamlyn Centre**), Université de Leeds (**National Facility for Innovative Robotic**

Systems), Université de Sheffield (**Sheffield Robotics**), Université de Liverpool (**Centre for Autonomous Systems Technology**), Université de Southampton (**Autonomous Systems USRG**), Université de Warwick (**Warwick Manufacturing Group**), Université de Bristol (**Bristol Robotics Laboratory**), Loughborough University, University of Oxford (**Centre for Doctoral Training in Autonomous Intelligent Machines and Systems**)

En 2015, plusieurs universités (de Cambridge, Edinbourg, Oxford, University College London et Warwick) ont également créé avec l'EPSRC, l'Alan Turing Institute, qui est le centre national pour la science des données, basé à la British Library. L'Institut a été nommé d'après le célèbre mathématicien et cryptologue britannique Alan Mathison Turing (1912-1954), qui a posé des fondements de l'informatique et contribué à la réflexion sur l'intelligence artificielle, en particulier avec le « test de Turing ».

Il existe également des structures spécifiques de partenariat public/privé dédiées à l'innovation, appelées Catapult Centres, gérées par le Technology Strategy Board (renommé en 2014 « Innovate UK » ¹). Créés à partir de 2012, ces centres facilitent la R&D grâce à une collaboration entre chercheurs académiques et entreprises. Parmi ces centres, Digital Catapult effectue notamment des recherches en intelligence artificielle.

¹ *Innovate UK est un organisme public britannique financé par le Department for Business, Energy & Industrial Strategy (BIES)*

L'Open Data Institute (ODI) figure également parmi les acteurs majeurs de la R&D en intelligence artificielle. Il s'agit d'une société privée à but non lucratif fondée en 2012 par Tim Berners-Lee (un des inventeurs du World Wide Web et actuel président du World Wide Web Consortium) et Nigel Shadbolt (Professeur en intelligence artificielle à l'University of Southampton). Le réseau de l'ODI comprend des entreprises, des établissements publics ainsi que des individus et vise à innover dans le domaine des données.

Ainsi, des liens entre les différents centres de recherche, qu'ils soient publics ou privés, ont été renforcés par une mise en réseau au niveau national.

Au titre de la structuration des activités dans le domaine, tant au niveau académique qu'industriel, on peut également citer The Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB) qui est la société savante principale en intelligence artificielle au Royaume-Uni. Créée en 1964, elle réunit des membres au-delà du territoire britannique et bénéficie ainsi d'une portée internationale.

Le financement de la R&D en intelligence artificielle : vers un renforcement des capacités nationales suite au Brexit ?

Les dépenses de R&D du Royaume-Uni (1,72 % du PIB en 2014) sont globalement inférieures à celles de la France (2,26 %) ou de l'Allemagne (2,84 %). La Confédération de l'Industrie Britannique (CBI) a ainsi recommandé une augmentation de la part des dépenses publiques de R&D afin d'atteindre un objectif de 3 % du PIB en 2025. Cependant, les budgets alloués à la recherche au Royaume-Uni bénéficient d'un excellent retour sur investissement : chaque livre sterling dépensée en crédit d'impôt à la R&D entraînerait un investissement complémentaire évalué entre 1,53 et 2,35 livres sterling au Royaume-Uni¹.

Une augmentation du financement de la R&D de 4,7 milliards de livres sur quatre ans, jusqu'en 2020-21, a été annoncée dans l'Autumn Statement 2016. Cet investissement ciblera notamment les domaines prioritaires que sont l'IA, la robotique, les technologies énergétiques intelligentes, les biotechnologies et la 5G. Ces fonds seront attribués de deux façons : à travers un nouveau fonds industriel « Industrial Strategy Challenge Fund » (géré par Innovate UK et par les Research Councils), et *via* l'attribution de financements par UK Research & Innovation (UKRI) qui regroupera les deux entités précédentes. Actuellement, le gouvernement britannique investit environ 6,3 milliards de livres par an pour la R&D – en plus du financement des établissements d'enseignement supérieur – ainsi que de 2 milliards de livres en crédits d'impôt pour les entreprises qui effectuent des activités de R&D.

¹ Source : <https://taj-strategie.fr/cout-chercheurs-monde/> 31 janvier 2017

Au niveau fiscal, le ministre anglais des finances George Osborne a annoncé en juillet 2017 – suite au Brexit – que l'impôt sur les sociétés continuerait de baisser, pour atteindre 15 % en 2020 contre 20 % en 2016. Ce changement impacterait indirectement le crédit d'impôt dédié à la recherche (*Research and Development Tax Credit*) en faisant augmenter sa valeur.

Cependant, des chercheurs universitaires s'inquiètent d'une éventuelle diminution des fonds de R&D suite au Brexit¹. En effet, l'Union européenne est active sur le thème de l'intelligence artificielle de multiples façons, par exemple à travers le financement de projets par le programme Horizon 2020 ou encore avec sa contribution financière à des événements tels que l'« *European Robotics Forum* » qui aura lieu en mars 2017 à Édimbourg. Plusieurs organismes européens existent dans le domaine de l'intelligence artificielle, notamment l'association EuRobotics AISBL et le partenariat public-privé pour la robotique en Europe (SPARC).

L'affirmation d'une stratégie britannique en intelligence artificielle, notamment dans un cadre industriel

Innovate UK a créé en 2013 un « Robotics and Autonomous Systems Special Interest Group » (RAS-SIG). Ce pôle d'intérêt commun est chargé d'analyser l'état de la recherche en robotique et systèmes autonomes ainsi que d'élaborer une stratégie nationale unissant des acteurs publics et privés. Le pôle RAS-SIG a publié en 2014 une stratégie pour la robotique, l'automatisation et les systèmes autonomes à l'horizon 2020².

La recherche académique en IA entre dans le cadre d'une stratégie nationale de soutien à l'industrie et en particulier à l'« Industrie 4.0 ». Une nouvelle stratégie industrielle, appelée « Building our Industrial Strategy: Green Paper » a été lancée le 23 janvier 2017. Cette stratégie se concentre autour de dix piliers, dont le premier est « Investir dans la science, la recherche et l'innovation ». Dès 2012, la robotique et les systèmes autonomes (RAS) avaient été identifiés par le gouvernement britannique comme étant l'un des huit domaines technologiques à développer dans le cadre de l'« UK Industrial Strategy ».

La chambre des communes (House of Commons) a publié un rapport sur la robotique et l'intelligence artificielle en octobre 2016³. Dans ce rapport, les députés demandent la création d'une Commission permanente sur l'intelligence artificielle. Cette nouvelle Commission serait basée à l'Alan Turing Institute, qui est le centre national pour la science des données. Le

¹ Nature, "[How scientists reacted to the Brexit](#)", 24 June 2016 ; Nature, "[Brexit watch: UK researchers scramble to save science](#)", 22 July 2016 ; The New York Times, "[Brexit May Hurt Britain Where It Thrives: Science and Research](#)", 17 October 2016

² RAS-SIG, [Robotics, Automation and Autonomous Systems \(RAAS\) 2020 Strategy](#), July 2014

³ House of Commons, Science and Technology Committee, [Robotics and artificial Intelligence, Fifth Report of Session 2016-17](#), October 2016

gouvernement a émis une réponse¹ prudente, en janvier 2017, affirmant qu'il continuerait à financer les Catapult Centres et le nouveau Industrial Strategy Challenge Fund (ISCF).

En novembre 2016, le Government Office for Science a publié un rapport² examinant les opportunités et implications de l'IA pour la société et pour le gouvernement. Ce rapport aborde les questions de la gestion des risques éthiques et juridiques ainsi que des avantages que pourrait apporter l'IA à la société et au gouvernement.

Les nécessaires adaptations du marché du travail et des formations universitaires

Les nouvelles technologies de robotique et d'IA seraient susceptibles à la fois de créer de nouveaux métiers et de faire disparaître des emplois. Le rapport « Technology at work: V2.0 »³ publié en janvier 2016 par l'entreprise financière Citi, en collaboration avec l'Université d'Oxford, prévoit que 35 % des emplois au Royaume-Uni pourraient être automatisés d'ici à 2020. Cela place néanmoins le pays dans la fourchette basse en termes de risques (57 % en moyenne dans l'OCDE). Ce rapport souligne également qu'une automatisation accrue pourrait conduire à de plus grandes inégalités économiques. Les emplois vulnérables sont principalement ceux peu qualifiés, tels que les travailleurs en centres d'appels, les vendeurs et les ouvriers des industries manufacturières. Inversement les emplois à faible risque d'automatisation sont ceux aux activités variées requérant de l'intelligence sociale et de la créativité. Les secteurs moins impactés seraient l'éducation, la santé et le social.

Des risques d'automatisation pèsent également sur des emplois qualifiés. Par exemple en finance, les algorithmes « robo-advice » permettent de recommander des produits d'épargne et de placement à un client de manière similaire à un conseiller financier⁴. Dans le secteur juridique, l'entreprise ROSS Intelligence a créé un service d'analyse de la jurisprudence à destination des cabinets d'avocats. La diffusion de ce type de « robots » pourrait conduire à des bouleversements sectoriels. Le cabinet de conseil Deloitte a publié en janvier 2016 des estimations sur l'impact de l'automatisation sur le marché du travail. Ainsi, ces estimations prédisent

¹ House of Commons, Science and Technology Committee, *Robotics and artificial Intelligence, Fifth Report of Session 2016 – [Appendix: Government response](#)*, 11 January 2017.

² Government Office for Science, [Artificial intelligence: opportunities and implications for the future of decision making](#), 9 November 2016

³ Oxford Martin School and Citi, [Technology at Work v2.0: The Future Is Not What It Used to Be](#), January 2016

⁴ John Gapper, ["Robots are better investors than people"](#), *Financial Times*, March 16, 2016

que 114 000 emplois du secteur juridique au Royaume-Uni pourraient être automatisés au cours des 20 prochaines années¹.

Ces transformations à venir du marché du travail interrogent sur la nécessité d'adapter les formations universitaires². Selon le site Whatuni, le Royaume-Uni compte au moins 36 universités qui offrent une formation en IA, représentant un total de plus de 117 diplômes universitaires³. L'intelligence artificielle pourrait également impacter la formation avec l'arrivée de nouveaux services numériques. Le professeur Rose Lucking de l'University College London (*Knowledge Lab, Institute of Education*) effectue des recherches sur les liens entre numérique et éducation, en particulier en ce qui concerne les applications l'IA pour l'éducation ainsi que l'apprentissage analytique et le *big data*.

Au niveau des débouchés, de nombreuses entreprises d'intelligence artificielle sont basées au Royaume-Uni⁴.

Quelques exemples d'innovations et d'applications marquantes en intelligence artificielle

- Pour la première fois dans l'histoire britannique un véhicule autonome a circulé dans les rues publiques, en octobre 2016. Ce véhicule autonome a été développé au Royaume-Uni par *Transport Systems Catapult* dans le cadre du projet *LUTZ Pathfinder*, initié en 2014.



Self-Driving Pods. Credits: Transport Systems Catapult

- Après avoir obtenu des autorités britanniques l'autorisation de lever les restrictions de vol, c'est depuis son *Prime Air Lab*, installé dans le *Cambridgeshire*, qu'*Amazon* a effectué sa première livraison par drone d'un colis à un client en décembre 2016.

- La société *DeepMind* spécialisée dans l'apprentissage profond (*deep learning*), créée par le Dr. Demis Hassabis, titulaire d'un doctorat de l'*University College London*, a été rachetée par *Google* en 2014 pour plus de 628 millions de dollars américains.

¹ Deloitte, "[Automation transforming UK industries](#)", Press Releases, 22 January 2016

² David Matthews, "[The robots are coming for the professionals](#) - Do universities need to rethink what they do and how they do it now that artificial intelligence is beginning to take over graduate-level roles?", *Times Higher Education*, July 28, 2016

³ Ces Bachelors' et Master's Degree comportent des intitulés avec les termes : "Artificial Intelligence", "Robotics", "Computer Science", "Intelligent Systems", "Automation" ou encore "Cybernetics". <https://www.whatuni.com/degree-courses/search?subject=artificial-intelligence>

⁴ Par exemple : *Business Insider UK*, "[10 British AI companies to look out for in 2016](#)", Jan. 5, 2016.

- La start-up de reconnaissance vocale *VocalIQ* développée par le Professeur Steve Young de *l'Information Engineering Division* à l'Université de Cambridge a été rachetée par *Apple* en 2015 et contribue à l'évolution du logiciel *Siri*.

- Le Dr. Zoubin Ghahramani, professeur en ingénierie d'information à l'Université de Cambridge et membre de l'Alan Turing Institute, est co-fondateur de la start-up *Geometric Intelligence* qui est basée à New York et qui développe des technologies de *machine learning*. L'entreprise de quinze personnes a été rachetée fin 2016 par Uber et deviendra *l'Uber AI Labs* à San Francisco.

5. En Suisse

Note réalisée avec le concours de l'Attaché de coopération scientifique et universitaire de l'Ambassade de France en Suisse

1. Une contribution scientifique française essentielle

Marvin Lee Minsky, de l'Université de Carnegie-Mellon, définit à la fin des années cinquante l'intelligence artificielle comme « *la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique* ». Ces dix dernières années, le domaine de l'intelligence artificielle a connu des progrès très importants notamment avec l'invention d'algorithmes d'apprentissage profond (*deep learning* algorithms, parfois aussi appelés réseaux de neurones) dus notamment au Français Yann LeCun¹. La stratégie mise en place dans ces algorithmes, mime le fonctionnement des neurones du cerveau où des connections se créent, disparaissent ou se renforcent en fonction des stimuli. Notons que dans le cadre biologique il n'y a pas d'**intelligence** sans **apprentissage**. Il s'agit à partir d'une phase d'apprentissage sur des objets connus, comme par exemple des images, des sons, des caractères, de pouvoir déterminer des paramètres d'un algorithme afin que celui-ci puisse donner des réponses pour des objets totalement nouveaux avec une erreur faible. Ici l'erreur est définie comme la différence entre la réponse de l'algorithme et la réponse exacte obtenue pour des données test. De manière plus précise, en déterminant un **nombre fini de paramètres** de l'algorithme avec un ensemble **d'apprentissage fini fixé**, on veut que l'erreur entre la réponse exacte et la réponse de l'algorithme soit minimale pour un ensemble **test fini fixé**. La détermination des paramètres se réalise avec un procédé d'optimisation. Un algorithme d'apprentissage profond comporte quatre éléments :

- des **données de spécification** (données d'entraînement et données de test) ;
- une **fonction coût** à minimiser pour déterminer les paramètres ;
- une **procédure d'optimisation** ;
- un **modèle** permettant de relier les variables d'entrée aux variables de sortie.

Donnons tout de suite quelques exemples que ces algorithmes permettent de traiter : la reconnaissance de visages, la traduction automatique, la reconnaissance d'objet dans une scène vidéo ou la classification (partition d'un ensemble en classes d'éléments de propriétés

¹ <https://www.college-de-france.fr/site/yann-lecun/>

similaires). Pour tous ces exemples, la dimension des « exemples d'apprentissage » doit être de l'ordre de 10 millions de cas et la dimension de l'ensemble des paramètres est facilement de 1 000 ou plus. Cette difficulté des dimensions est appelée la « malédiction de la grande dimension » (*the curse dimensionality*). Ainsi, pour être efficace, les algorithmes d'apprentissage profond doivent utiliser des **données massives** et disposer d'une bonne puissance de calcul. De plus, plusieurs couches d'algorithmes d'apprentissage profond associées à de la classification sont nécessaires pour traiter les exemples complexes évoqués plus haut. Cela correspond à une décomposition en tâches plus simples pour traiter un problème compliqué. Dans le cas de la reconnaissance d'un visage, on ne cherchera pas à reconnaître la totalité du visage en une seule fois, mais certains de ses éléments séquentiellement. Ce point de vue de l'apprentissage conjugué au traitement séquentiel de tâches plus simples et l'augmentation de la puissance des calculateurs ont permis à l'intelligence artificielle de sortir de l'impasse du traitement global de tâches complexes.

La « pensée magique » est inappropriée

L'approche consistant à créer d'énormes ensembles de données, puis à appliquer les algorithmes d'apprentissage profond dont on dispose aujourd'hui est inappropriée pour 'créer de la connaissance' même si elle est souvent proposée. En effet, il n'existe pas d'algorithme universel (pouvant traiter toutes sortes de données), cela peut se démontrer mathématiquement. De même, suivant les problèmes que l'on veut traiter, il faut introduire de l'*a priori* sur la connaissance recherchée (ce qui en termes techniques de la discipline s'appelle les hyperparamètres et les termes de pénalisation).

Il faut quand même citer la réussite de la stratégie d'apprentissage combinée à une mesure de similarité (principes fondamentaux des algorithmes d'apprentissage profond) mise en œuvre pour le traducteur automatique « Google translate ». À partir de bases de données de textes traduits, Google translate propose une traduction par similarité à l'existant déjà traduit. Cet algorithme d'apprentissage profond donne de bons résultats là où des approches analytiques plus linguistiques avaient échoué.

Des questions théoriques à approfondir

Il n'y a pas d'algorithme universel possible

Du point de vue de la logique, il n'est pas possible d'avoir un processus d'inférence général (c'est-à-dire valable pour n'importe quel ensemble) à partir d'un nombre limité d'exemples. Cette contradiction logique est levée pour les algorithmes d'apprentissage car ils ne fournissent qu'une réponse probabiliste. Néanmoins même dans le cadre probabiliste, si on recherche une grande généralité pour le domaine d'application, c'est-à-dire qu'un algorithme puisse s'appliquer pour toutes les données la réponse reste négative. Prenons l'exemple des algorithmes de classification par apprentissage, il ne sera pas possible d'en trouver un qui est meilleur que les autres (ce résultat est connu mathématiquement sous le nom du « *no free*

lunch theorem ») si il doit s'appliquer à toutes les données possibles. Dans la pratique pour avoir un taux d'erreur faible on introduit de la connaissance *a priori* sur les données que nous traitons.

*La malédiction de la grande dimension*¹

Outre la question de la puissance de calcul que le traitement de grands ensembles de données pose, il y a d'autres aspects théoriques limitant.

La minimisation de la fonction coût dépendant d'un très grand nombre de variables et non convexes est mathématiquement difficile car ces fonctions peuvent avoir des comportements beaucoup plus complexes qu'en petites dimensions. On peut s'en faire une représentation en pensant à un paysage de montagne vu de haut et très accidenté qui présente beaucoup de sommets et de creux, sachant que l'on recherche le creux le plus profond, celui-ci ne sera pas facile à déterminer.

Ce résultat est pénalisant car le calcul des paramètres nécessite toujours la minimisation d'une fonction coût.

Dans les ensembles de grande dimension les **éléments ne sont pas uniformément répartis**, de la même manière que dans le ciel que l'on observe, les étoiles de la Voie lactée, ne sont pas distribués de manière homogène. Cela veut dire que l'on doit prendre en compte cette difficulté lorsque l'on recherchera des « structures » (ou de la connaissance) dans ces grands ensembles de données.

2. Des réussites et des potentialités

L'exemple de « Google translate » a déjà été cité, mais on pourrait aussi mentionner la publicité ciblée sur le web, l'analyse de données issues du génome, la reconnaissance de la parole comme domaine d'activité où les algorithmes d'apprentissage profond ont permis une réelle avancée.

Parmi les potentialités des algorithmes d'apprentissage profond, on peut noter :

- la médecine personnalisée ;
- les humanités digitales ;
- le déplacement autonome des robots ou des véhicules ;
- les outils pédagogiques pour l'apprentissage des sciences.

¹ <https://www.college-de-france.fr/site/yann-lecun/seminar-2016-02-19-15h30.htm>

3. L'intelligence artificielle en Suisse

Il faut tout d'abord mentionner le centre Google Research Europe localisé à Zurich qui sera dirigé par le Français Emmanuel Mogenet. Le groupe se concentrera sur trois axes principaux de recherche: outre l'intelligence artificielle, il se consacrera aussi à la compréhension et au traitement automatique du langage naturel utilisée notamment pour les applications vocales, ainsi qu'à la perception artificielle.

Le SGAICO¹ - Swiss Group for Artificial Intelligence and Cognitive Science. Le SGAICO rassemble des chercheurs, des praticiens, ainsi que d'autres personnes intéressés au domaine de l'intelligence artificielle et des sciences cognitives (AI/CO), sous la houlette de la société suisse d'informatique. Le SGAICO vise la promotion des technologies intelligentes pour l'innovation dans la société. Il apporte une plateforme d'échanges en AI/CO où l'industrie et les universités peuvent se rencontrer. Jana Koehler² est présidente du SGAICO.

Il y a aussi le laboratoire d'intelligence artificielle et d'apprentissage automatique³ dirigé par le professeur Boi Faltings à l'École polytechnique fédérale de Lausanne, ainsi que l'Automatic Control Laboratory⁴ à l'École polytechnique fédérale de Zurich.

L'institut de recherche Idiap⁵ (anciennement Institut d'intelligence artificielle perceptive), situé à Martigny (État du Valais, Suisse), est une fondation de recherche autonome, indépendante et à but non lucratif spécialisée dans la gestion d'informations multimédia et dans les interactions homme-machine multimodales. Fondé en 1991 par la municipalité de Martigny, l'École polytechnique fédérale de Lausanne, l'Université de Genève et Swisscom, il est reconnu internationalement pour ses travaux en traitement et reconnaissance de la parole, apprentissage artificiel, vision par ordinateur, authentification biométrique et interface homme-machine. Il est associé à l'EPFL par un plan de développement commun. Son budget, qui s'élève à plus de 10 millions de francs suisses, est financé à 50 % par des projets de recherche récompensés selon des processus concurrentiels et à 50 % par des fonds publics. Alors qu'il employait une trentaine de personnes en 2001, l'Idiap a en 2016 une centaine d'employés, dont 80 chercheurs.

L'Idiap, alors dirigé par le professeur Hervé Boulard⁶, a été l'institution hôte du Pôle de recherche national IM2 (2001-2013) dédié à la Gestion interactive et multimodale de systèmes d'information.

¹ <http://www.s-i.ch/fr/sgaico/>

² <https://user.enterpriselab.ch/~takoehle/>

³ <http://ic.epfl.ch/intelligence-artificielle-et-apprentissage-automatique>

⁴ <https://control.ee.ethz.ch/overview.en.html>

⁵ <https://www.idiap.ch/linstitut>

⁶ <http://people.idiap.ch/boulard> *boulard [at] idiap.ch*

La propriété des données en Suisse

La loi suisse sur la protection des données (LPD) est une des plus restrictives d'Europe, mais son domaine d'application est restreint. À ce propos, notons qu'un certain nombre de sociétés pensent installer en Suisse leur centre de données pour cette raison.

Pour la gestion des données de médecine personnelle la LPD n'est pas adaptée. En effet, à partir d'applications dites de « bien-être », comme les montres connectées ou des objets connectés, les entreprises qui développent les applications informatiques enregistrent des données physiologiques personnelles (rythme cardiaque, pression artérielle...) sur leurs serveurs. Ces données personnelles sont par la suite utilisées par ces entreprises comme elles le souhaitent. Les milieux scientifiques suisses sont conscients des dérives que cela pourrait entraîner, et développent des solutions innovantes où l'individu décide de l'usage de ses données de santé. Citons par exemple la plateforme midata¹ développée par la Haute école spécialisée de Berne et qui accueille déjà deux projets, dont un sur la sclérose en plaque avec des patients de l'hôpital universitaire de Zurich.

Du point de vue légal en Suisse, les juristes considèrent que le droit ne se préoccupe que des états et non des processus. Ainsi des principes généraux suffisent et il n'y a pas besoin de lois spécifiques pour encadrer ces technologies.

4. Un regard critique sur l'intelligence artificielle

Dans son ouvrage *La société automatique* le philosophe Bernard Stiegler voit dans le développement de l'intelligence artificielle et des automatismes un risque de disparition de l'abstraction et de la pensée. L'argument repose sur l'antagonisme entre l'automatisme et le processus cognitif reposant sur l'observation, la construction d'un modèle abstrait permettant d'interpréter l'observation et l'analyse de l'adéquation des deux. L'intelligence artificielle a permis de développer des procédures de « saisie prédictive » de texte qui utilisent les textes que l'on a déjà écrits et la fréquence d'apparition d'une séquence. Ce qui est le plus utilisé se substitue à ce que l'on pourrait créer ou penser.

¹ <https://www.midata.coop/fr.html>

ANNEXE 3 : SUR LE DROIT COMPARÉ DE LA ROBOTIQUE

Les éléments suivants sont issus des deux ouvrages, *Droit des robots* de MM. Alain et JérémY Bensoussan et *Comparative handbook : robotics technologies law*, ainsi que du livre blanc du Symop *Droit de la robotique* publié en 2016. Ils font souvent l'objet d'un parti-pris favorable à la création d'une personnalité juridique pour les robots.

En 2015, le marché de la robotique représentait plus de 26 milliards de dollars (contre 15 milliards de dollars en 2010). La robotique et l'intelligence artificielle pourraient représenter une révolution industrielle d'ampleur au moins comparable à celle d'Internet. **Le développement de technologies d'intelligence artificielle peut potentiellement concerner quatre secteurs d'investissement clés : la défense, l'industrie, l'entreprise et le secteur personnel.** La Commission européenne a évalué, en 2012, que la seule robotique de service pourrait constituer un marché de plus de 100 milliards d'euros à l'horizon 2020. La robotique pose des questions juridiques essentielles. Le soutien des pouvoirs publics est important, au travers de partenariats public-privé, à l'instar du plan France Robots Initiatives (100 millions d'euros), et de la mise en place d'une fiscalité attractive afin de favoriser les investissements dans le domaine de la robotique¹.

Il est proposé, ci-dessous, d'effectuer des **comparaisons juridiques sur les points suivants : le statut légal du robot, les régimes de responsabilité appliqués aux robots, la protection des données collectées et stockées par le robot, la propriété intellectuelle de la création robotique, la reconnaissance des contrats concernant ou conclus par des robots, les robots dans le secteur de la santé et le cas des voitures autonomes.**

1. Le statut légal du robot

La rupture technologique provoquée par les progrès importants de la robotique et de l'intelligence artificielle pourrait poser la question de la création d'une personnalité juridique pour les robots. La création d'une telle personnalité juridique permettrait d'offrir à l'entité robot dotée d'intelligence et de liberté décisionnelle une catégorisation juridique alignée sur ses capacités, ainsi qu'une insertion apaisée et sécurisée dans le tissu social. L'élément clé dans la conception d'une personnalité robot est la liberté décisionnelle que celui-ci aura acquis grâce à son intelligence et à sa capacité d'apprentissage. Bien que le robot voie ses capacités limitées par l'humain qui les lui aura attribuées, il demeure qu'il jouira d'une certaine

¹ Cf. par exemple le dispositif d'amortissement accéléré prévu par la loi de finances pour 2014.

liberté, même résiduelle, son comportement ne pouvant être complètement prédit.

Tableau comparatif sur le statut légal du robot

Afrique du Sud	S'il n'existe pas de statut légal ni de législation propre aux robots, certaines dispositions peuvent leur être appliquées, comme le <i>Protection of Personal Information Act 4</i> de 2013, ou le <i>Medecines and Related Substances Amendment Act 14</i> de 2015. De même, le <i>South African Bureau of Standards (SABS)</i> est un organe autonome établissant les normes de standardisation qui, de fait, concernent les robots, notamment en appliquant les standards ISO 10218-1:2011 et ISO 8373:2012.
Allemagne	Il n'existe pas de statut légal du robot ; cependant, de nombreuses dispositions contenues notamment dans le BGB (code civil), dans le StGB (code pénal), dans le TMG (loi sur les télémedias), ou encore le BDSG (loi fédérale sur la protection des données).
Belgique	Il n'existe pas de statut légal du robot ; le robot est reconnu comme un objet juridique mais est défini d'un point de vue légal comme une machine industrielle.
Brésil	Il n'existe pas de statut légal du robot ou de la machine intelligente, à l'exception des drones qui sont reconnus comme des objets juridiques. Les robots demeurent des objets juridiques non identifiés d'un point de vue légal.
Chine	Les robots n'ont pas de statut légal reconnu ; cependant, ils sont identifiés comme des « produits », et sont donc considérés comme des objets conformément à au droit des biens chinois (<i>PRC Property law</i>).
Costa Rica	Les robots sont classés dans la catégorie des biens, et sont considérés comme des objets juridiques.
Espagne	Il n'existe aucun droit des robots ni statut légal des robots ; les robots sont considérés comme des objets, et n'ont donc pas de personnalité juridique. Aucun projet de loi concernant l'octroi d'un statut légal du robot n'est pour l'instant à l'étude.
États-Unis	Bien que les technologies robotiques représentent un marché émergent considérable, la législation tant fédérale que locale n'accorde pas de statut légal aux robots.
France	Les robots n'ont pas de statut légal, et sont considérés comme des biens. Cependant, la personnalité juridique étant une construction juridique qui n'est pas exclusive aux êtres humains, les robots pourraient bénéficier d'une adaptation de la législation en se voyant attribuer une personnalité juridique.
Grèce	Il n'existe pas de législation propre aux robots ; cependant, certains textes visent la régulation de certains types de technologies robotiques, comme la législation sur les drones ou la législation sur les voitures autonomes.

Israël	Il n'existe pas de législation spécifique aux robots ; cependant, de nombreuses lois peuvent s'appliquer aux robots, comme sur la propriété, la propriété intellectuelle ou la responsabilité relative aux dommages causés par une opération robotique.
Italie	Si la directive 2006/42/CE (directive Machines) a été transposée dans l'ordre juridique interne, il n'existe pas de statut légal du robot. Certaines dispositions, notamment en matière de responsabilité, existent, mais elles ont été pensées pour des machines industrielles, et non pour des robots.
Japon	La législation nationale apporte plusieurs définitions contenant le terme robot - comme celle introduite par le cadre juridique de la radio qui définit le robot comme une machine ou un équipement sans-fil qui transmet automatiquement des données. Il n'existe cependant pas de législation pertinente pour les robots, du fait de la difficulté d'appliquer une régulation spécifique pour chaque type de robot.
Liban	Il n'existe aucun statut légal du robot ; une réflexion peut cependant être menée sur l'octroi d'une personnalité juridique aux robots.
Portugal	Les robots n'ont pas de statut légal ; ils sont considérés par la loi comme des « machines » et sont en ce sens des objets soumis aux lois et réglementations en application, notamment en matière de commerce, de transports, d'investissement ou de production de machines et d'équipements.
Royaume-Uni	Il n'existe pas de cadre juridique applicable aux robots. Cependant, la réflexion autour de l'élaboration d'une personnalité juridique robot pourrait notamment s'appuyer sur les sept « principes de la robotiques » proposés par l'EPSRC et l'AHRC ou par la stratégie « RAS 2020 » proposée par le groupe d'intérêt RAS-SIG à destination du Gouvernement.
Suisse	Il n'existe aucun statut légal du robot, et seule la loi sur la transparence peut pour le moment s'appliquer aux robots.

2. Régimes de responsabilité

Afrique du Sud	Un régime de responsabilité spécial est appliqué aux robots : celui de la responsabilité du fait des produits défectueux en vertu de la loi CPA (<i>Consumer Protection Act 68 of 2008</i>). Ainsi, la section 61 de la loi CPA impose une responsabilité objective aux fabricants, importateurs, distributeurs et vendeurs de biens, notamment de robots.
Allemagne	Le régime de responsabilité allemand distingue principalement : le régime de responsabilité contractuelle du régime de responsabilité civile, et la responsabilité objective de la responsabilité subjective. Aucun de ces régimes ne semble convenir pour un régime de responsabilité du robot, étant donné qu'il peut être très complexe – voire impossible – de retracer la

	<p>chaîne de causalité et l'intentionnalité ou la négligence dans un cas de dommage causé par un robot autonome à une personne juridiquement responsable. Cependant, dans l'état actuel de la législation, le régime de responsabilité objective semble le plus adéquat à appliquer aux robots ; et un régime de responsabilité subjectif s'applique au fabricant en cas de comportement intentionnel ou négligeant entraînant une défaillance du robot.</p>
Belgique	<p>Le régime de responsabilité est principalement subjectif ; mais les mécanismes de responsabilité actuels ne sont pas nécessairement adaptés à la coopération homme-robot : étant donné que les robots sont des ensembles complexes mêlant matériel informatique et logiciels, parfois intégrant de l'intelligence artificielle, il peut être impossible, ou très long et très coûteux, de prouver une faute qui aurait causé des dommages. Il en serait de même dans la perspective de l'application d'un régime de responsabilité objective.</p>
Brésil	<p>Il n'existe pas de dispositions spécifiques concernant le régime de responsabilité appliqué aux robots ; de manière générale, le régime de responsabilité objectif est appliqué dans les relations civiles ou commerciales, étant donné que les robots sont considérés comme des objets.</p>
Chine	<p>Le régime de responsabilité par défaut est le régime objectif, et le régime de responsabilité subjectif ne s'applique que dans certains cas avec des circonstances spécifiques. Les robots étant considérés comme des produits, leur régime de responsabilité est défini par la loi sur la qualité des produits (<i>RPC Product Quality Law</i>). Ainsi, le fabricant et le vendeur assument une responsabilité objective concernant la qualité du produit délivré au client. De même, en cas de dommages physiques ou à la propriété du client, le vendeur assume une responsabilité civile fondée sur la faute et doit donc, selon l'article 42 de la loi sur la qualité des produits, assurer la compensation du dommage causé.</p>
Costa Rica	<p>Les régimes de responsabilité objective et de responsabilité subjective sont tous deux appliqués en fonction des cas ; de manière générale, quiconque cause un dommage à autrui a une obligation de le réparer. En fonction de la tâche exécutée par le robot et du risque qui y est associé, l'un de ces deux régimes s'applique. Cependant, ces deux mécanismes ne sont pas adaptés à la coopération humain-robot.</p>
Espagne	<p>Un régime de responsabilité objective s'applique de manière générale s'applique aux produits ; cependant, dans le cas très techniques, la Cour peut inverser la charge de la preuve et imposer au fabricant de démontrer que le produit n'est pas défectueux. Un robot étant un produit très complexe dont il peut être difficile de prouver la défaillance, ces précédents d'inversement de la charge de la preuve dans des cas très techniques peuvent s'appliquer pour des dossiers liés à des dommages causés par des robots.</p>

États-Unis	Il est probable que de nombreux types d'équipements robotiques ou quasi-robotiques seront soumis à des standards de responsabilité stricte du fait des produits défectueux, notamment dans le cas des robots médicaux. Il serait en effet difficile de prouver la négligence du fabricant dans le cas des robots. En outre, le droit des contrats fait émerger d'autres formes de responsabilités ; ainsi, les fabricants et les distributeurs doivent veiller à inclure des renonciations de garanties appropriées et des limitations de responsabilité dans leur documentation sur les produits, les divulgations et les contrats pour la vente de ces produits.
France	Le robot est considéré comme une chose dans le droit civil français. La responsabilité du fait des choses, en tant que régime de responsabilité objectif prévu par l'article 1384 du code civil, signifie que pour qu'elle soit appliquée, la chose doit être impliquée dans le dommage et qu'elle joue un rôle actif (comme le fait d'être en mouvement ou de toucher la victime) dans l'occurrence dudit dommage. C'est l'individu considéré comme « gardien » de la chose qui est responsable de la réparation du dommage cause. Cependant, si le dommage est causé par une faille de sécurité du robot, le régime de responsabilité pour le dommage causé par la chose s'applique au fabricant du robot.
Grèce	Le fabricant assume une responsabilité du fait des produits défectueux : en vertu de l'article 6 de la loi 2251/1994, si un produit défectueux cause un quelconque dommage à un consommateur ou à leur propriété, le fabricant doit compenser ce dommage indépendamment du fait qu'il y ait eu négligence ou faute de sa part. Dans le cas des robots, identifier la partie responsable en cas de dommage peut être un exercice complexe, d'autant plus qu'une technologie robotique peut être composée d'un logiciel en <i>open source</i> qui, par nature, vise à être modifié. Selon le professeur David Vladeck (Université de Georgetown), une distinction d'application de régime de responsabilité selon le caractère semi-autonome ou pleinement autonome d'un robot. Dans le cas d'un robot semi-autonome, donc servant d'outils à un humain ou à une entité juridique, il est possible d'appliquer la responsabilité du fait des produits défectueux, assumée par le fabricant. Cependant, dans le cas de robots pleinement autonomes, il est plus adéquat de recourir à un régime de responsabilité objective quand il est presque impossible d'identifier la partie fautive, ou si c'est la machine elle-même qui est fautive, étant donné que celle-ci ne dispose pas de capacité juridique.
Israël	En l'absence de régime de responsabilité spécifique aux robots, différents régimes de responsabilité pourraient s'appliquer selon les cas : la loi sur la responsabilité civile des produits défectueux de 1980 (<i>Defective products Law</i>), qui impose un régime de responsabilité objective sur les fabricants et importateurs de produits défectueux lorsqu'une partie ayant subi le dommage soit physiquement blessée. L'ordonnance de responsabilité civile (<i>Torts Ordinance</i>) indique que, pour attester d'une négligence, le plaignant doit prouver l'existence d'une obligation de diligence, le manquement à ce devoir par le défendeur, et une connexion causale entre le manquement à ce devoir et le dommage causé au plaignant. Les termes d'un contrat entre un fabricant et un utilisateur peuvent soulever une cause d'action fondée sur la garantie fabricant/importateur du robot.

Italie	<p>En vertu de la directive 59/92/EC, le produit est supposé comme sûr par définition. Ainsi, en cas de dommage causé par l'utilisation du produit, il est de la responsabilité du fabricant que de vendre uniquement des produits sûrs.</p> <p>Trois parties prenantes peuvent être responsables d'un dommage commis par un robot : le fabricant, ce qui implique que sont engagées la responsabilité du créateur et celle du programmeur ; le distributeur, si le bien a été produit et importé depuis l'extérieur de l'Union européenne ; et l'utilisateur, qui peut être responsable de dommages causés à l'encontre de tierces parties par un objet sous son contrôle et durant son utilisation.</p>
Japon	<p>La loi du 1^{er} juillet 1995 sur la responsabilité civile des produits défectueux dispose que le fabricant est considéré comme responsable et doit compenser le dommage subi ; cependant, le fabricant ne saurait être responsables des dommages résultant uniquement de la qualité du produit en lui-même.</p>
Liban	<p>La loi de 2000 sur la protection des consommateurs tient une personne pour responsable si cette personne a introduit un produit défectueux sur le marché. Le cas du régime de responsabilité des robots en cas de dommage causé à une personne tierce soulève de nombreuses questions : est-ce qu'un robot est responsable dans ses actions et ses révisions ? Qui doit être tenu responsable ?</p>
Portugal	<p>Concernant le régime de responsabilité contractuelle, il n'existe aucune règle applicable aux accords dans le champ de la robotique ou impliquant des robots. Concernant le régime de responsabilité non contractuelle, il est nécessaire de distinguer trois cas :</p> <ol style="list-style-type: none"> 1. Le cas où la partie subissant le dommage doit prouver la faute, ce qui correspond à la règle générale du droit Portugal pour le régime de responsabilité non contractuelle. 2. Le cas où la faute est présumée en raison du non-respect du devoir de chacun de surveiller tout objet déplaçable en leur possession et le devoir de chacun à prévenir le danger qui résulte d'une dangereuse activité déclenchée ou bénéficiant à cette personne ou légale entité. 3. Le cas où le régime de responsabilité objective s'applique, lorsque le fabricant met le produit sur le marché ou quiconque ayant le contrôle sur un véhicule terrestre. <p>En vertu du décret-loi 383/89, le régime de responsabilité d'un fabricant dont le produit a causé un dommage à cause d'un défaut de fabrication. Dans ce cas-ci, les robots peuvent à la fois qualifiés de « fabricants » (robots concepteurs) et de « produits ».</p>

Royaume-Uni	<p>Les robots étant considérés par le droit britannique comme des produits, ils se voient appliqués le régime de responsabilité du fait des produits défectueux (<i>product liability</i>). Cependant, le cadre juridique actuel n'est pas adapté pour le cas des robots, car il ne peut pas apporter toutes les solutions aux problèmes de responsabilité qui émergent et émergeront de l'utilisation des robots. La principale difficulté réside dans la complexité du robot en tant que produit composé de éléments matériels et de logiciels et dont la création et réalisation d'opérations est difficilement traçable.</p> <p>Les robots n'ayant pas de personnalité juridique, le régime de responsabilité du fait des produits défectueux ou des actes qui causent des dommages ne pourra reposer sur le robot lui-même. Cependant, désigner la personne responsable pour les dommages causés par un robot sera une tâche difficile, du fait du nombre de sujets impliqués par la création, la commercialisation et le fonctionnement des robots.</p>
--------------------	---

3. Protection des données collectées et stockées par le robot

Afrique du Sud	<p>La personne morale ou physique qui détermine pourquoi et comment le robot collecte les informations personnelles est responsable de la protection de ces informations.</p>
Allemagne	<p>Le propriétaire du robot est responsable du traitement des données et de son respect aux lois de protection des données personnelles en vigueur. Le règlement 2016/679 sur la protection des données personnelles sera, en outre, applicable dès 2018.</p> <p>Lorsque des données personnelles sont stockées ou traitées par le robot, le propriétaire ou l'opérateur traitant les données doivent en notifier l'autorité fédérale de protection des données personnelles. En outre, chaque Lander dispose, en plus de l'autorité fédérale, d'une autorité de protection des données indépendante.</p> <p>Les responsables privés de robots doivent enregistrer ces procédures auprès des autorités de supervision compétentes.</p> <p>L'enregistrement obligatoire ne s'applique pas si le responsable du traitement a désigné un responsable de la protection des données ou si le responsable du traitement recueille, traite ou utilise des données à caractère personnel. Ainsi, la plupart des entreprises allemandes ont nommé un responsable de la protection des données.</p> <p>Si une partie tierce procède au traitement des données personnelles sous l'autorité de l'entité responsable, la loi considère que cette partie tierce fait partie de l'entité responsable et le transfert ou l'accès aux données personnelles ne requiert pas d'autorisation légale.</p> <p>La section 9 de loi fédérale sur la protection des données (BDSG) pose des exigences fondamentales en matière de mesures organisationnelles et techniques de sécurité des données.</p>

Belgique	<p>Le code pénal belge permet à un propriétaire de poursuivre un individu ayant piraté ou saboté les données contenues dans un robot, qui est considéré comme un système informatique.</p> <p>Les données collectées et stockées par un robot sont protégées par la loi du 8 décembre 1992 sur la protection de la vie privée, cette loi transposant la directive 9546/EC du 24 octobre 1995.</p> <p>Tout propriétaire de robot doit notifier en amont toute opération de traitement des données collectées et stockées par le robot à la commission vie privée</p> <p>Le règlement général sur la protection des données adopté en 2016 par le Parlement européen impose une protection des données par défaut pour tous les produits, services et systèmes utilisant des données personnelles.</p> <p>Le propriétaire du robot doit assurer la sécurité et la confidentialité de ses données personnelles collectées et stockées par le robot.</p>
Brésil	<p>Il n'existe pas de régime juridique de protection des données personnelles spécifique aux robots ; de fait, les dispositions générales s'appliquent, notamment l'article 5 de la Constitution et les dispositions que l'on peut trouver dans le code civil, dans le code pénal, dans le code de la consommation ou encore dans la loi civile de l'Internet du 23 avril 2014.</p>
Chine	<p>Si la loi sur la protection des informations personnelles (<i>Personal Information Protection Act</i>) adoptée en décembre 2015, elle ne répond pas aux « canons » européens de la protection des données : cette loi définit (et protège) les « informations » personnelles comme celles qui permettent d'identifier un utilisateur (nom, âge...) et les informations qui permettent la localisation et l'heure de consommation d'un service par un utilisateur.</p> <p>Cependant, il existe quelques textes, comme la décision du comité permanent de l'Assemblée nationale populaire sur le renforcement de la protection de l'information sur les réseaux, qui peuvent s'appliquer aux robots collectant et stockant des données à caractère personnel.</p>
Costa Rica	<p>Le propriétaire d'un robot gère la collecte et le stockage de données personnelles ; et les propriétaires ainsi que les vendeurs doivent respecter les dispositions de la loi n° 8968 sur la protection des individus concernant le traitement des données personnelles.</p>
Espagne	<p>La protection des données personnelles collectées par un robot ou un système informatique est assurée par la loi organique 15/1999 et le décret royal 1720/2007. Le propriétaire ou l'utilisateur d'un robot (désigné comme responsable du robot) doit respecter la législation de protection des données concernant le traitement de données personnelles au travers d'un robot. De même, le responsable gérant le système de traitement des données doit respecter certaines obligations posées par l'Agence espagnole de protection des données (AEPD), notamment sur la notification du type de données détenues, l'identité du responsable, ou encore le type de traitement de données et son but. Le responsable du robot est également tenu d'adopter les mesures techniques nécessaires pour assurer la sécurité des données personnelles afin d'empêcher leur altération, leur perte, ou tout traitement ou accès non autorisé.</p>

États-Unis	<p>Il est possible d'attribuer à certaines données une paternité, ce qui les place sous la protection du droit d'auteur. C'est notamment le cas des lignes de code qui composent l'intelligence du robot. Les nombreux composants constituant un équipement robotique peuvent être également protégés par les lois sur le secret industriel. La question de la commercialisation, de l'utilisation d'un robot et des opérations et traitements effectués celui-ci pose la question du consentement de personnes tierces dont les données personnelles peuvent être collectées (par exemple, une personne photographiée par un drone).</p>
France	<p>Les activités robotiques sont soumises au respect de la loi Informatique et libertés de 1978. Ainsi, les propriétaires de robots gérant le système de traitement des données doivent respecter les obligations posées par la CNIL, c'est-à-dire la notification standard sur les utilisations du robot, sur le type de logiciel utilisé, sur les systèmes de sécurité installés pour protéger les données d'intrusions tierces non autorisées, sur les données personnelles stockées dans le robot, et sur les informations fournies aux utilisateurs concernant le traitement de leurs données personnelles. Tout traitement de données personnelles doit être signalé en amont à la CNIL, et l'utilisation de données « sensibles », comme les données médicales, doit être autorisée par la CNIL. En outre, les propriétaires de systèmes de traitement de données personnelles sont soumis à une obligation de sécurité et de confidentialité des données.</p>
Grèce	<p>La collecte et le traitement de données à caractère personnel par un robot est régi par la loi 2472/1997 (DPL) et la loi 3471/2006 (PECL) qui transposent respectivement les directives européennes 95/46/EC et 2002/58/EC. En vertu de la loi DPL, le responsable des données (<i>data controller</i>) doit notifier par écrit à l'Autorité hellénique de protection des données (DPA) l'établissement d'une base de données ou le début d'un traitement de données. En vertu de la loi DPL, le traitement de données personnelles doit être confidentiel. Il doit être effectué exclusivement par les personnes agissant sous l'autorité et les instructions du responsable des données. Les informations stockées par un robot mais également les informations concernant la structure, les caractéristiques et les opérations effectuées par un robot peuvent être, dans certaines circonstances, être qualifiées d'informations confidentielles. Cependant, la législation ne fait pas mention spécifiquement des informations confidentielles stockées dans un robot ; néanmoins, le cadre juridique actuel peut s'appliquer aux robots.</p>
Israël	<p>Les utilisateurs de robots sont soumis aux mêmes lois de la protection des données qu'une personne collectant des données d'une différente manière.</p> <p>La loi 5741-1981 sur la protection de la privée indique que les détenteurs de bases de données contenant des informations appartenant à plus de 10 000 personnes ont l'obligation de s'enregistrer et d'identifier l'objectif guidant l'entretien d'une telle base de données, les moyens de collecte des données contenues dans cette base, et les mesures de sécurité utilisées pour protéger les données.</p>

<p>Italie</p>	<p>Le code italien de protection des données régit toutes les activités de collecte de données à caractère personnel. Ainsi, toute personne dont les données personnelles sont connectées doit être informée de l'identité de celui la personne physique ou morale qui les recueille et dans quel but. Le fait que les robots ne bénéficient pas d'une personnalité juridique a un impact significatif sur ce le cas des opérations de collecte de données effectuées par un robot : qui est la personne qui doit informer la personne dont les données sont collectées ?</p> <p>Le droit italien a posé des exigences minimales de sécurité, la section 32 du code de protection des données indiquant notamment que les données doivent être protégées par de mesures qui seraient adéquates et cohérentes avec l'état de l'art de la technologie. Ainsi, si les mesures minimums de sécurité ne peuvent être évitées, la loi introduit un niveau de standard bien supérieur en matière de sécurité des données, car ces standards doivent être cohérents avec l'état de la technologie mais également du niveau de risque et du type de données traitées.</p>
<p>Japon</p>	<p>Étant donné qu'un robot est en capacité de stocker et de traiter des données personnelles, l'utilisateur d'un tel robot peut être considéré comme un opérateur commercial détenant des informations personnelles, qui sont protégée par la loi de protection des informations personnelles, texte amendé en 2015 dont l'application est prévue en 2017.</p> <p>Si un robot collecte plus de 5 000 dossiers d'informations personnelles en répondant à des visées commerciales, l'utilisateur du robot est un opérateur commercial qui doit répondre à de nombreuses exigences, telles que la spécification obligatoire de l'utilisation des informations personnelles ou la mise en place des mesures nécessaires permettant d'assurer une gestion sécurisée des informations personnelles. À noter qu'avec l'amendement du texte, à partir de 2017, tout robot stockant et traitant des données sera considéré comme un opérateur commercial, même si ce robot a stocké et traité moins de 5 000 dossiers d'informations personnelles.</p>
<p>Liban</p>	<p>Le Liban ne dispose pas d'autorité nationale de protection des données ni de loi de protection des données.</p>
<p>Portugal</p>	<p>Si le robot est capable de collecter et de traiter des données personnelles, le responsable des données sur le territoire portugais, ou utilisant un équipement situé sur le territoire portugais, doit se conformer aux règles applicables comme énoncées dans la loi n°67/98 sur la protection des données personnelles. Une notification de traitement des données personnelles par le responsable des données doit être envoyée en amont à l'Autorité portugaise de protection des données.</p> <p>Dans le cas du traitement des données « sensibles », leur traitement est interdit sauf si une autorisation spéciale est accordée par l'autorité ou si une des quelques exceptions juridiques liées aux raisons de l'intérêt public s'appliquent. Le responsable des données est soumis à un devoir de confidentialité et doit installer des mesures techniques et organisationnelles adéquates afin de protéger les données personnelles contre les accidents, dommages ou actes malveillants.</p>

Royaume-Uni	<p>Le traitement de données personnelles est encadré par la loi sur la protection des données de 1998, qui dispose que le respect de la législation sur le traitement des données personnelles incombe au responsable de la base de données. Le responsable doit notifier en amont l'<i>Information Commissioner's Office</i> (ICO) de tout traitement de données personnelles en vertu de la section 18 de la loi sur la protection des données.</p> <p>Le responsable de la base de données doit également respecter de nombreuses obligations imposées par la loi, notamment le traitement équitable et licite des données uniquement aux fins légales spécifiées pour lesquelles les données ont été obtenues et la mise en place de mesures de sécurité afin de prévenir tout traitement non autorisé ou illicite, ainsi la perte, la destruction ou l'endommagement des données.</p>
Suisse	<p>Les provisions juridiques standards contenues dans la loi fédérale sur la protection des données et les différentes lois cantonales de protection des données sont appliquées aux questions de collecte et de traitement de données à caractère personnel par des robots.</p>

4. Propriété intellectuelle de la création robotique

La production et l'utilisation de robots et de technologies similaires dans le domaine public, industriel ou privé, peut être limitée par divers mécanismes de protection de la propriété intellectuelle, notamment avec l'existence de brevets qui protègent l'innovation technique, le droit d'auteur ou le savoir-faire, non breveté mais dont le détournement peut engager la responsabilité de l'auteur dudit détournement.

Afrique du Sud	<p><u>Droit d'auteur</u> : le droit d'auteur est actuellement défini par les dispositions de la loi sur le droit d'auteur de 1978 (<i>Copyright Act 98 of 1978</i>). Le droit d'auteur n'a pas besoin d'être enregistré ; il s'applique automatiquement. Cependant, la création robotique n'est pas prévue par ce texte. Néanmoins, un <i>Copyright Amendment Bill</i> devrait être promulgué au cours de l'année 2017 qui pourrait affecter le droit d'auteur pour les créations robotiques.</p> <p><u>Dépôt de marque</u> : si un robot peut constituer une marque déposée, il n'est pas prévu qu'une création robotique puisse déposer une marque.</p> <p><u>Brevet</u> : si un modèle de robot peut être breveté, il n'est pas prévu qu'une création robotique puisse être brevetée.</p>
Allemagne	<p>Dans l'état actuel de la législation, il n'est pas possible pour un robot de détenir une quelconque propriété intellectuelle ou industrielle, car ils sont considérés au regard de la loi comme des objets, et ne sont donc pas porteurs de droits. Seuls les humains peuvent exercer un droit de propriété sur une création.</p>

Belgique	Ni le droit national ni le droit communautaire ne prévoient la création par un robot ou un ordinateur. Aucun texte ne considère les robots comme des fournisseurs de service ; les robots sont considérés comme des objets, des outils, et de fait, cela exclut leurs créations du champ de la protection de la propriété intellectuelle et du droit d'auteur tel que défini par le droit communautaire.
Brésil	Les brevets, les marques déposées et le régime de droits d'auteur peuvent offrir aux robots une protection intellectuelle. Cependant, la loi sur la propriété industrielle et la loi du 19 février 1998 sur le droit d'auteur ne permettent pas aux robots de breveter leurs créations ni d'être considérés comme des auteurs. Ces lois considèrent que la création ne peut être qu'une humaine, et non issue de machines.
Chine	<p><u>Brevet</u> : les robots étant considérés au regard de la loi comme des produits, il est difficile qu'ils puissent déposer un brevet. Cependant, si une création humaine résultant de l'usage d'un robot ou une création proprement robotique remplissent les critères de brevetage, il est possible d'obtenir une protection de la propriété intellectuelle en vertu du droit des brevets.</p> <p><u>Droit d'auteur</u> : en vertu de la loi sur le droit d'auteur (<i>PRC Copyright Law</i>), une création souhaitant bénéficier du droit d'auteur doit remplir trois critères : l'originalité, la duplication et la légalité. Or, le critère d'originalité est ici conditionné au fait qu'une création doit être accomplie par la pensée indépendante d'un auteur et doit refléter la philosophie et l'intelligence de jugement de l'auteur, plutôt qu'une simple copie ou compilation de matériaux. De fait, les robots ne peuvent satisfaire le critère d'originalité, et leurs créations ne peuvent donc pas être protégées en vertu du droit d'auteur.</p> <p><u>Dépôt de marque</u> : comme de nombreux pays, la Chine a ratifié la Classification de Nice sur les biens et les services, dont la classe 7 catégorise les robots comme des biens et dont les actions en tant que fournisseur de service ne sont pas explicitement pris en compte. Cependant, à l'exception des robots répondant à la définition de la classe 7, les robots peuvent être classés selon leur fonction et leur tâche. Ainsi, les robots médicaux, d'entraînement et d'éducation peuvent être considérés comme de classe 10, 41, 44, etc.</p> <p><u>Savoir-faire</u> : le droit de la concurrence (<i>PRC Anti-Unfair Competition Law</i>) protège le secret d'affaires ; à ce titre, le propriétaire légal du robot peut protéger les techniques du robot en tant que secret d'affaires.</p>
Costa Rica	Dans l'état actuel de la législation, un robot ne peut détenir un droit d'auteur ou des droits intellectuels ni déposer une marque sur une création qu'il aurait développée.
Espagne	Si les robots, en tant que produits, peuvent être brevetés et protégés par le droit d'auteur et le dépôt de marque, la création robotique ne peut en bénéficier.
États-Unis	<p>De nombreuses classifications de brevet de conception et de brevet d'utilité peuvent s'appliquer aux robots.</p> <p>Les marques déposées pour la robotique sont protégées de la même manière que les autres outils électroniques ou techniques.</p> <p>La législation ne fait aucune mention de la protection de la propriété intellectuelle ou industrielle de la création robotique.</p>

<p>France</p>	<p><u>Brevet</u>: il est possible d'accorder un brevet à une création résultant d'un processus industriel impliquant un ordinateur ou un robot.</p> <p><u>Droit d'auteur</u>: le régime juridique actuel réserve la paternité d'une création à un individu, excluant de fait une paternité exclusivement robotique.</p> <p><u>Savoir-faire</u>: si les créations produites par des composants robotiques ne sont pas éligible à la protection assurée par le code de la propriété intellectuelle, le savoir-faire, défini comme « un ensemble d'informations pratiques non brevetées résultant de l'expérience et de l'expérimentation », peut être utilisé comme un outil pour protéger la création robotique.</p> <p><u>Dépôt de marque</u>: de même que dans les pays ayant ratifié la Classification de Nice, les robots sont considérés comme des biens et leurs actions en tant que fournisseur de services ne sont, de fait, pas prises en compte.</p>
<p>Grèce</p>	<p><u>Droit d'auteur</u>: toute création effectuée au recours d'un robot en tant qu'outils est protégé par la loi 2121/1993 sur le droit d'auteur. Cependant, il n'est pas possible de protéger le droit d'auteur d'une personne non humaine, autant une entité juridique qu'un robot.</p> <p><u>Dépôt de marque</u>: la Grèce a ratifié la Classification de Nice, et les robots sont à ce titre couverts par la classe 7 ; mais au vu du nombre de services que peut offrir un robot, il est également possible de les inclure dans d'autres classes, comme la classe 12, 28, 41, 44, etc.</p> <p><u>Brevet</u>: un robot peut, en tant que tel ou en tant que nouvelle méthode, faire l'objet d'un brevet. Cependant, la législation ne prévoit pas de dispositions permettant le brevetage de la création robotique.</p>
<p>Israël</p>	<p>Les robots peuvent être, en tant que le résultat d'efforts collaboratifs, être protégés par les droits de la propriété intellectuelle ; cependant, le droit d'auteur n'a nul besoin d'enregistrement et s'applique automatiquement.</p>
<p>Italie</p>	<p>Comme dans le reste de l'espace communautaire, la création de robots peut bénéficier d'une protection de la propriété intellectuelle, et les robots peuvent être brevetés. Le droit d'auteur (<i>copyright</i>) n'existe pas en tant que tel en Italie ; le régime de protection de la propriété intellectuelle protège les inventions et les droits des auteurs, ces droits étant une conséquence naturelle de la création d'un bien protégé. Cette protection est accordée dès lors que le produit est marqué d'un caractère de nouveauté d'originalité. Le savoir-faire est quant à lui protégé par le secret industriel.</p>
<p>Japon</p>	<p><u>Brevet</u>: la loi sur le brevet indique qu'une invention industriellement applicable peut être autorisée à obtenir un brevet. Les robots peuvent être protégés en tant qu'inventions. Une invention est brevetée lorsqu'elle exprime une nouveauté, une démarche inventive et a une application industrielle. Un inventeur ne peut être qu'un être humain, les robots n'étant pas des sujets titulaires de droits et d'obligations, et ne peuvent de fait pas exprimer leurs idées.</p> <p><u>Droit d'auteur</u>: si le robot peut assister l'humain dans le travail de création, un travail est considéré comme une production dans laquelle les idées et pensées sont exprimées de manière créative et qui relève de n'importe quel champ artistique. De fait, un robot ne peut être considéré comme un auteur et en recevoir la protection.</p>

Liban	Il existe de nombreux systèmes de protection de la propriété intellectuelle, notamment le brevet, le droit d'auteur ou le dépôt de marque et de concept, qui peuvent être utilisés et combinés pour offrir une protection totale à toute invention, et notamment aux robots.
Portugal	<p><u>Brevet</u> : toute nouvelle invention dans le domaine de la robotique qui implique une activité inventive et appropriée à une utilisation industrielle peut être brevetée.</p> <p><u>Droit d'auteur</u> : il existe une protection juridique du droit d'auteur pour les programmes informatiques en tant qu'ouvrage littéraire ; il n'existe aucune raison pour laquelle cela ne pourrait être appliqué aux logiciels robots.</p> <p><u>Dépôt de marque</u> : le dépôt de marque peut être utilisé dans le domaine de la robotique afin d'identifier le résultat de développement du robot (le produit final, le logo et le nom du robot peuvent être protégés).</p>
Royaume-Uni	<p><u>Brevet</u> : en vertu de la loi sur les brevets de 1977, une invention est brevetable si elle est nouvelle, si elle implique une démarche inventive, si elle offre des applications industrielles et n'est pas interdit par la loi, ce qui permet aux robots satisfaisant ces critères d'être brevetés.</p> <p><u>Conception</u> : la conception d'un modèle de robot peut remplir les conditions pour être protégée en vertu de la loi sur les modèles enregistrés de 1949 et la loi sur le droit d'auteur, les modèles et le brevet de 1988.</p> <p><u>Droit d'auteur</u> : en vertu de la loi sur le droit d'auteur, les modèles et le brevet de 1988, un logiciel peut être protégé par le droit d'auteur en tant qu'ouvrage littéraire, et tout comme une compilation de données si cela représente un travail original et non copié qui a nécessité de la part de l'auteur l'application de compétences, du discernement et du travail. Si une compilation de données peut correspondre à la définition de la base de données telle qu'énoncée dans la section 3A (1) de cette loi, alors les bases de données contenues dans les robots peuvent être protégés.</p> <p><u>Dépôt de marque</u> : les robots étant considérés aux yeux de la loi comme des produits, les fabricants peuvent avoir recours au dépôt de marque. Ainsi, les robots et les services délivrés par celui-ci peuvent être protégés par le dépôt de marque.</p> <p><u>Création robotique</u> : il n'existe aucune loi ni aucune jurisprudence traitant de la question de l'application des droits de propriété intellectuelle aux créations robotiques.</p>

5. Contrats concernant ou conclus par des robots

Afrique du Sud	Il est possible de conclure un contrat de manière électronique, comme dans de nombreux pays mais la question des contrats intelligents (passés par et entre un ou plusieurs robots) n'est pas abordée. Cependant, un contrat peut techniquement être conclu par une personne utilisant un robot, il est de fait nécessaire de développer concernant la capacité de contractualisation d'un robot.
-----------------------	---

Allemagne	Les logiciels ne peuvent exprimer de déclaration d'intention. Cependant, des contrats conclus par des robots peuvent être possible au regard du droit allemand si et seulement si ces déclarations sont associées à un utilisateur et sont considérés comme appartenant à l'utilisateur. Un contrat conclu par des robots peut également être valide si les parties contractantes ont décidé, du fait de leur liberté contractuelle, d'une procédure où ce sont les robots qui établissent les clauses du contrat. Néanmoins, ces solutions ne permettent pas de résoudre tous les problèmes liés à la question des contrats conclus par des robots.
Belgique	Les contrats conclus automatiquement par des robots ne sont pas abordés par le droit national. Cependant, les deux parties d'un contrat doivent stipuler dans celui-ci qu'un système autonome pourra prendre un certain nombre de décisions qui les contraindront.
Brésil	Les contrats conclus par des robots ou impliquant une décision effectuée par un robot ne sont pas abordés ou régulés par un cadre juridique spécifique ; cependant, les décisions automatiques relèvent de la responsabilité de ceux qui utilisent ces systèmes autonomes ou qui décident d'incorporer ce système dans le processus contractuel.
Chine	Un robot, considéré comme un produit, peut également être l'objet d'un contrat de vente entre deux parties. De fait, le robot étant un produit, il ne peut être reconnu comme l'une des parties concluant un contrat mais peut en être l'objet.
Costa Rica	Il n'existe aucun cas enregistré de contrat passé automatiquement par un robot.
Espagne	Les contrats passés par des robots ne sont pas abordés par la législation espagnole. De fait, des contrats civils ou commerciaux passés par l'intermédiaire d'interactions robotiques peuvent se voir appliquées les règles contractuelles classiques sur les ventes et les droits et obligations généraux prévus par le code civil et le code de commerce.
États-Unis	De nombreuses activités, surtout financières, ont recours, grâce aux <i>Fintech</i> , à des contrats passés automatiquement par des robots. Ainsi, il est admis que 70 % des opérations à Wall Street sont effectuées par des robots.
France	Le robot peut être l'objet d'un contrat ; cependant, les contrats « intelligents », passés par des robots notamment <i>via</i> une chaîne de blocs utilisant une monnaie virtuelle comme le <i>Bitcoin</i> (qui n'est pas reconnue comme monnaie dans le droit interne) ne sont pas reconnus par le code civil.
Grèce	Les systèmes autonomes, comme les robots, sont de plus en plus utilisés dans les transactions, et notamment dans le cadre de transactions boursières. À ce stade d'installation et d'utilisation de contrats intelligents, il est possible d'affirmer que les concepts et principes traditionnels de droit civil, en identifiant la plateforme opératrice derrière le robot.

Italie	<p>Bien que de nombreuses entreprises y aient recours, les contrats concernant la vente ou la location de robots ne disposent pas encore de modèles ou de standards comme c'est le cas des contrats concernant les ordinateurs. Le système juridique italien est un système de responsabilité limitée : les dommages doivent être prouvés, et la charge de la preuve est du ressort du plaignant. Les dommages doivent être la conséquence directe et immédiate de actes ou faits préjudiciables, et les dommages doivent être évités de façon prévisible. La jurisprudence montre que, de manière répétée, lorsque les produits « high-tech » sont concernés, il existe un écart de connaissance entre le fabricant ou le vendeur et l'acheteur, ce qui signifie que le fabricant a une meilleure connaissance des implications que peut avoir l'utilisation de certaines technologies, notamment les robots. Le fabricant/vendeur doit respecter un certain niveau de diligence. Ainsi, il est essentiel que l'acheteur soit dûment informé sur de nombreux aspects du robot.</p>
Japon	<p>Il n'existe pas de standards ou modèles pour les contrats concernant les robots. Du fait de la complexité des structures et des mécanismes composant les robots, il est difficile pour les parties d'un contrat impliquant un robot de déterminer si un robot rempli les exigences des utilisateurs et acheteurs. De nombreux différends concernant les contrats sur les robots naissent du fait des interprétations des différentes parties du contrat. Les contrats conclus par des robots, ou automatiquement, sont d'usage courant au Japon dans le secteur services financiers à destination des particuliers. Cependant, les robots contractants participent à la procédure de formation du contrat, mais la décision finale demeure prise par un humain.</p>
Liban	<p>La vente et l'achat de robots peuvent être régis par deux textes : le code des obligations et des contrats, et le code de la consommation. L'article 401 du code des obligations et des contrats impose notamment deux obligations aux vendeurs : une obligation de délivrer l'objet de la vente, et une obligation de garantir une utilisation paisible de l'objet par l'utilisateur et de le garantir contre les défauts. La vente de robot devrait être conditionnée au respect de mesures de sécurité et d'exigences techniques ; cependant, cela n'est pas spécifié pour le moment dans la législation. Il n'existe pas non plus d'encadrement des contrats robots par le code des obligations et des contrats, ce texte ne limitant la liberté de contrat qu'au respect de l'ordre public, des règles obligatoires et des valeurs éthiques et morales.</p>
Portugal	<p>Il n'existe pas de cadre légal spécifique aux contrats sur les robots ; le principe de liberté contractuelle des parties s'applique. Il n'existe aucune législation spécifique concernant les contrats intelligents ; cependant, les contrats passés <i>via</i> des ordinateurs sont des extensions de la volonté humaine, en ce sens que la machine agit en fonction des instructions et scénarios calculés et saisis par le programmeur.</p>
Royaume-Uni	<p>Un contrat juridiquement exécutoire exige, entre autres, l'intention de créer des relations juridiques et, si le contrat n'est pas conclu directement par une partie, le pouvoir d'agir en son nom. Lorsque des moyens électroniques et des machines automatisées sont utilisés comme des outils qui facilitent simplement la conclusion d'un contrat conformément à des exigences strictement prédéterminées établies par les parties, alors un accord contraignant peut être formé. Cependant, si des robots peuvent, grâce à un haut niveau d'automatisation, conclure des contrats, il demeure un flou important. Les robots ne sont dès lors plus considérés comme des outils permettant la communication des intentions de chaque partie mais un</p>

	système autonome capable de décisions immédiate et en interaction avec son environnement, et de fait les parties peuvent ne pas être informées sur le contenu ou l'existence même du contrat et donc de la création de relations juridiques. Il existe un flou juridique sur la question des contrats passés par, tout particulièrement lorsque ces contrats contiennent des clauses contraignantes.
--	--

6. Les robots dans le secteur de la santé

Il existe une certaine diversité des robots médicaux, notamment les robots d'aide au diagnostic, les robots chirurgicaux, et les robots d'aide aux soins ou de maintien de l'autonomie des personnes âgées ou handicapées.

Les robots médicaux sont considérés par les agences nationales et internationales de produits de santé comme des "dispositifs médicaux" ; l'autorité nationale compétente en France est l'Agence nationale de sécurité du médicament et des produits de santé (ANSM).

La mise en circulation sur le marché des robots médicaux est encadrée par des exigences de conformité issues de la réglementation européenne transposée dans le code de la santé publique, et sont classés en quatre catégories, chacune d'elle étant associée à des règles d'évaluation et de contrôle spécifiques.

Le code de la santé publique impose aux professionnels de santé d'engager leur responsabilité en cas d'incident liée à l'utilisation d'un robot médical défaillant ; en outre, il existe trois normes ISO encadrant la sécurité des robots médicaux : la norme ISO 13482:2014 révisée par la norme ISO/DIS 13485.2 qui énonce les exigences de sécurité pour les robots de soins personnel, et la norme ISO 13485:2003 qui fixe des exigences relatives à la qualité des dispositifs et services médicaux fournis.

La protection des données médicales en France est assurée par plusieurs textes. D'une part, la notion de données « relatives à la santé » sont définies par l'article 1111-7 et 1111-8 du code de santé publique, les délibérations de la CNIL et la décision du 6 novembre 2003 de la Cour de justice de l'Union européenne. D'autre part, la collecte et le traitement des données dites « sensibles », qui comprennent les données de santé telles que, est strictement encadré par la loi Informatique et libertés qui impose une obligation de déclaration des traitements de ces données personnelles à la CNIL.

Tableau comparatif

<p>Afrique du Sud</p>	<p>Les robots médicaux sont sujets de la loi de 1965 sur les médicaments et substances relatives, texte qui fut amendé par la loi de 2015 sur les médicaments et les substances relatives.</p> <p>L’Autorité de régulation des produits de santé d’Afrique du Sud (SAHPRA) régule l’utilisation des robots dans le secteur médical.</p>
<p>Allemagne</p>	<p>Les robots médicaux, désignés « dispositifs médicaux », sont notamment soumis aux dispositions de la loi sur les dispositifs médicaux (MGP) qui transpose les directives 93/42/CEE, 90/385/CEE et 98/79/CE. La régulation sur les dispositifs médicaux est, de fait, presque harmonisée au sein de l’espace communautaire. La commercialisation des dispositifs médicaux est conditionnée à l’obtention d’un marquage de conformité CE. La loi MGP ne contient aucune disposition concernant la responsabilité des dispositifs médicaux en cas de dommages causés. De fait, le régime général de responsabilité s’applique (objectif/subjectif ; contractuel/civil). La loi BDSG (section 3) définit les données médicales comme des données « sensibles », et leur traitement et utilisation est ainsi soumise à des exigences élevées. Le patient concerné doit en outre avoir déclaré expressément son consentement.</p>
<p>Belgique</p>	<p>La commercialisation de dispositifs médicaux est soumise aux dispositions de la directive 93/42 qui a été transposée dans l’arrêté royal du 15 juillet 1997 relatif aux dispositifs médicaux implantables actifs et dans l’arrêté royal du 18 mars 1999 relatif aux dispositifs médicaux. Ainsi, la commercialisation d’un dispositif médical est sujette à l’obtention préalable du marquage CE qui indique sa conformité avec les exigences de sécurité et de santé.</p> <p>L’arrêté royal du 18 mars 1999 oblige les professionnels de santé à avertir l’Agence fédérale des médicaments et des produits de santé de tout incident impliquant l’utilisation d’un dispositif médical ; ils peuvent être tenus responsables en cas de non-communication de tout incident.</p> <p>Plusieurs standards ISO sont appliqués aux dispositifs médicaux : ISO 13482:2014, ISO 13485:2003 et ISO/DIS 13485.2.</p>
<p>Brésil</p>	<p>L’Agence nationale de vigilance sanitaire (ANVISA) a édicté des règles spécifiques concernant l’usage de technologies dans le domaine médical, notamment sur les équipements et les produits, les logiciels, les exigences de sécurité, les données des patients, les exigences de gestion des documents électroniques, les technologies de vidéo et de surveillance et les applications de diagnostic et les produits liés.</p> <p>Bien qu’il n’existe pas de droit de la protection des données, le secteur médical est soumis à de nombreuses dispositions mentionnant le caractère sensible des données et enregistrements médicaux qui sont donc soumis à la confidentialité.</p>

Chine	<p>Le droit chinois tient une posture prudente à l'égard de l'usage de robots médicaux, tout en encourageant le développement. Le droit chinois prévoit ainsi une prise en charge des événements indésirables impliquant que le fabricant ou l'utilisateur d'un robot médical doit en faire rapport à la structure compétente. De même, sont prévus un système de rappel de produits défectueux et de sanctions à l'égard des entités ne respectant les dispositions précédentes.</p>
Costa Rica	<p>Les robots médicaux, dénommés « dispositifs médicaux », sont soumis aux dispositions du règlement n° 34482-S pour l'enregistrement, la classification, l'importation et le contrôle d'équipement et de matériel biomédical.</p> <p>Les dispositifs médicaux ne peuvent être utilisés sans respecter les exigences de sécurité et de santé des patients, des utilisateurs et des parties tierces. Ces exigences incluent notamment d'avoir un mandataire en justice au Costa Rica, un certificat de conformité, une description du matériel utilisé, et l'intégralité des documents fournis en langue étrangère traduits en espagnol et dûment formalisés par le consulat du Costa Rica.</p>
Espagne	<p>Les robots médicaux sont désignés comme des dispositifs médicaux. Le décret royal 1591/2009 vise à garantir la qualité des dispositifs médicaux en prévoyant des inspections et contrôles périodiques au moyen de différentes procédures. Ce même décret permet à l'Agence espagnole des médicaments et produits sanitaires d'autoriser de manière individuelle et spécifique selon des considérations relatives aux services de santé la commercialisation et le déploiement d'un dispositif médical qui n'a pas satisfait l'ensemble des procédures de validation indiquées à l'article 13 du décret, comme la vérification de certification CE, la déclaration de conformité CE ou l'obtention du label CE. Les fabricants de dispositifs médicaux sont soumis au régime de responsabilité objectif, et les professionnels de santé utilisant des dispositifs médicaux sont également soumis au régime de responsabilité civile.</p>
France	<p>L'Agence nationale de sécurité du médicament et des produits de santé considère les robots médicaux comme des dispositifs médicaux ; l'ANSM est en charge de la régulation du marché des systèmes médicaux.</p> <p>La commercialisation de dispositifs médicaux est soumise aux obligations de conformité indiquées dans le code de la santé publique et est conditionnée à l'obtention d'un marquage CE reflétant sa conformité aux exigences de sécurité et de santé précisées par la législation européenne. Les dispositifs médicaux sont classés en quatre catégories selon leur niveau de risque potentiel pour la santé.</p> <p>Les professionnels de santé sont dans l'obligation de reporter auprès de l'ANSM toute défaillance, tout incident ou tout risque d'incident causé par un dispositif médical. Si un incident impliquant l'usage d'un dispositif médical défectueux, leur responsabilité est mise en cause.</p> <p>Plusieurs standards ISO sont appliqués aux dispositifs médicaux : ISO 13482:2014, ISO 13485:2003 et ISO/DIS 13485.2.</p>

Grèce	<p>La Grèce a transposé les directives 98/79/EC et 93/42/CEE sur les dispositifs médicaux, impliquant que, comme tout État membre, la Grèce doit prendre les mesures nécessaires pour assurer la mise sur le marché et la mise en service des dispositifs médicaux uniquement si ceux-ci ne compromettent pas la sécurité et la santé des patients, utilisateurs ou d'autres personnes.</p> <p>La collecte et le traitement de données « sensibles », telles que les données médicales, est interdit par la loi DPL. Cependant, la collection et le traitement de données sensibles peuvent être permis par la DPA selon des conditions spécifiques, comme le consentement écrit de la personne à qui appartiennent ces données.</p>
Israël	<p>La Knesset a promulgué la loi 5772-2012 sur les dispositifs médicaux puis les règlements 5773-2013 sur l'enregistrement des dispositifs médicaux. L'ambition de ces textes est de réguler le champ des dispositifs médicaux en Israël, et notamment leur conception, invention, expérimentation, production, utilisation, distribution, exportation et importation afin de maximiser la protection des utilisateurs.</p>
Italie	<p>La directive 2007/47/EC a été transposée en Italie avec le décret législative¹ n°60 du 20 janvier 2010, qui a amendé le décret législative N° 47 du 24 février 1997. Selon le décret législatif n°60, un dispositif médical doit être développé et produit de manière à ce qu'il ne puisse, dans son usage ordinaire et prévu, mettre en danger la santé ou la sécurité du patient. Si un dispositif est installé et utilisé correctement mais représente néanmoins un risque pour le patient, le ministère de la santé peut ordonner son retrait et en interdire la vente et l'utilisation en informant la Commission européenne. De plus, si un utilisateur constate qu'un dispositif a causé un incident ou a un problème technique, il doit en informer le ministère afin que celui-ci en informe le fabricant, et évalue l'incident et les mesures à prendre.</p> <p>Les sections 75 à 94 du code en matière de la protection des données personnelles traitent de la question des robots médicaux. La section 78 dispose notamment que le médecin doit expliquer clairement au patient la manière dont ses données personnelles seront traitées pour la poursuite de son traitement. Le patient est amené à indiquer son consentement par écrit. En outre, les données personnelles de santé doivent être dûment protégées, et doivent notamment être cryptées avant toute communication par voie électronique.</p>

¹ Un « decreto legislativo » correspond, dans l'ordre juridique italien, à une ordonnance.

Japon	<p>En vertu de la loi sur les médicaments et les dispositifs médicaux de 2014, la production et l'utilisation de robots médicaux doivent respecter certaines exigences :</p> <ul style="list-style-type: none"> - L'enregistrement des entreprises de fabrication (et de leurs usines) auprès du ministère de la santé, du travail et des affaires sociales. - L'obtention d'une autorisation de fabrication et commercialisation du ministère - Pour être commercialisé au Japon, tout dispositif médical doit être approuvé en amont par le ministère. <p>Concernant la question de la protection des données médicales, il n'existe pour le moment aucune loi spécifique. En principe, la protection des données médicales est assurée par les dispositions de la loi sur la protection des informations personnelles.</p>
Portugal	<p>En vertu de la section 7 de la loi sur la protection des données personnelles, toute donnée médicale collectée, stockée ou traitée par un robot est considérée comme une donnée personnelle sensible.</p> <p>Il n'existe pas de dispositions spécifiques au régime de responsabilité pour les dommages causés par un dispositif médical. Dans ces cas, les règles générales sur le régime de responsabilité non contractuelle et le régime de responsabilité du fait des produits défectueux.</p>
Royaume-Uni	<p>L'utilisation de robots dans le secteur de la santé au Royaume-Uni est conditionnée par le règlement sur les dispositifs médicaux de 2002, qui transpose les directives 93/42/CEE, 98/78/CE et 90/385/CEE. Les fabricants de dispositifs médicaux doivent engager une procédure d'évaluation de la conformité de leurs produits et obtenir un marquage « CE » avant toute publicité ou vente. Ils doivent en outre être enregistrés auprès de l'Agence de régulation des produits médicamenteux et de santé (MHRA).</p>

7. Voitures autonomes

Trois aspects juridiques majeurs liés à l'exploitation des voitures autonomes peuvent être abordés :

- **La responsabilité liée aux voitures** : dans les États de Floride, du Nevada et du district de Columbia, la loi impute la responsabilité en cas de défaillance de la technologie autonome au fabricant du système autonome, que ce fabricant soit le constructeur du véhicule ou non. En France, le régime en vigueur impose l'indemnisation en cas d'accident au conducteur, ce qui implique une adaptation de la législation et/ou l'adoption d'un cadre réglementaire pour les véhicules autonomes ; dans l'attente d'un tel cadre, ce sont les contrats signés entre les parties prenantes qui répartiront la charge de la dette éventuelle de réparation.

- La **propriété des données techniques** : la question de la propriété des données recueillies et traitées est essentielle mais n'est pour l'instant pas départagée, impliquant que des décisions au cas par cas, l'octroi de licences d'utilisations de ces données et d'accords entre les parties prenantes seront nécessaires.

- La **protection des données à caractère personnel** : la directive STI impose aux États membres de veiller à ce que les données identifiantes soient protégées contre toute utilisation abusive, encourageant également l'utilisation de données anonymes et restreignant l'usage des données personnelles aux traitements nécessaires au bon fonctionnement des applications et services de systèmes de transport intelligent. Ainsi, la production et le déploiement des véhicules intelligents devront suivre, compte tenu des dispositions de la loi Informatique et libertés mais également du règlement européen encadrant la protection des données personnelles en Europe, une approche respectueuse de la protection des données à caractère personnel (*privacy by design*).

Tableau comparatif

Afrique du Sud	La législation sud-africaine n'a pas encore traité le cas de la voiture autonome.
Allemagne	L'Allemagne ayant ratifié la Convention de Vienne, seuls les véhicules dotés de systèmes partiellement autonomes (comme les systèmes d'aide à la conduite) sont autorisés à la circulation sur la voie publique, et le chauffeur doit avoir le contrôle du véhicule. Dans l'état actuel du droit de la circulation, le conducteur est strictement responsable des dommages causés. De fait, le conducteur est également responsable des dommages causés résultant d'une défaillance du système autonome du véhicule. En outre, le régime de responsabilité du constructeur réside dans la constatation de trois erreurs : une erreur de construction, une erreur de fabrication, et une erreur d'instruction.
Belgique	Il n'existe aucune règle spécifique concernant l'usage de véhicules sans chauffeur. Le référentiel juridique est la Convention de Vienne, qui a été transposée dans le code de la route belge en 1968, qui interdit donc la circulation de véhicules sans chauffeur sur la voie publique. Cependant, le gouvernement fédéral a approuvé en septembre 2016 un « code de bonnes pratiques d'expérimentation » des voitures autonomes dans des lieux public.
Brésil	La législation actuelle ne contient aucune disposition concernant les véhicules autonomes. Le Brésil a ratifié la Convention de Vienne sur la circulation routière, ce qui implique que le code de la route brésilien ne contient aucune disposition autorisant les véhicules autonomes. Cependant, l'expérimentation de voitures autonomes sur la voie publique n'est pas expressément interdite.

Chine	<p>Le gouvernement chinois soutient activement le développement de la R&D et l'expérimentation de la voiture intelligente. Cependant, l'usage de véhicules autonomes sur la voie publique n'est pas encore autorisé, et la responsabilité de toute infraction et dommage causé par un véhicule intelligent incombe, s'il n'y a pas de personne dans l'habitacle, à l'opérateur contrôlant le système du véhicule.</p>
Costa Rica	<p>Le code de la route du Costa Rica ne contient aucune disposition autorisant l'utilisation de véhicules sans chauffeur, ni de police d'assurance spécifique pour ces véhicules. Dans l'état actuel de la législation sur la propriété intellectuelle, les données techniques générées par les voitures intelligentes seront la propriété du constructeur automobile et du fournisseur de composants.</p>
Espagne	<p>Bien qu'aucune législation relative à l'utilisation de véhicules sans chauffeur n'ait été adoptée, la Direction générale du trafic a adopté l'instruction 15/V-113 en 2015 permettant l'expérimentation de prototypes de véhicules autonomes sur des routes ouvertes au public si certains critères sont remplis.</p>
États-Unis	<p>Les constructeurs ont la possibilité d'expérimenter leurs véhicules sans chauffeur sur des routes ouvertes au public après avoir obtenu en amont une permission officielle. La Californie, le district de Columbia, la Floride, le Michigan et le Nevada ont déjà adopté des lois autorisant l'expérimentation de véhicules sur route. Les lois de ces États exigent cependant la présence d'un conducteur expérimenté à l'intérieur du véhicule.</p> <p>Afin de répondre aux défis de confidentialité et de sécurité soulevés par la collecte d'un grand nombre de données, les constructeurs automobiles ont développé les « Principes de protection de la vie privée des consommateurs pour les véhicules, les technologies et les services » qui déterminent la collecte et l'usage des informations des utilisateurs des véhicules connectés. Sept principes sont mis en avant : transparence ; choix ; le respect du contexte ; la « minimisation » de la collecte de données, la dé-identification et la rétention légitime ; la sécurité des données ; l'intégrité et l'accessibilité ; et la responsabilité.</p> <p>Afin de minimiser les risques liés à la circulation des véhicules intelligents, les constructeurs ont incorporé des principes de sécurité au cœur du processus de conception (<i>safety by design</i>), et ont établi en 2015 l'Auto-ISAC qui permet de faciliter les échanges d'importants volumes d'informations en temps réel sur les faiblesses éventuelles des dispositifs autonomes et des contremesures envisagées. En outre, les constructeurs se sont engagés de longue date aux côtés de nombreuses tierces parties telles que des technologues de la sécurité, des organisations non lucratives, des programmes gouvernementaux et des groupes de travail, des universités et des initiatives de STEM afin de développer technologies et usages de sécurité pour les véhicules.</p>

<p>France</p>	<p>L'article R. 412-6-I du code de la route indique que tout véhicule en mouvement doit avoir un conducteur. Cependant, la loi du 17 août 2015 dite « loi de transition énergétique pour la croissance verte » contient des dispositions introduisant le cadre expérimental visant à promouvoir l'expérimentation et le déploiement de véhicules propres, incluant les voitures sans chauffeur.</p> <p>Dans l'état actuel de la législation, le conducteur du véhicule est responsable en cas d'accident de la route. Cependant, cette législation ne peut être appliquée telle quelle aux accidents causés par des véhicules autonomes car le conducteur n'a pas le contrôle direct du véhicule.</p> <p>La loi relative à la transition énergétique pour la croissance verte du 17 août 2015 habilite le Gouvernement à agir par ordonnance concernant l'autorisation d'expérimentation de véhicules semi-autonomes ou autonomes sur la voie publique. En ce sens, l'ordonnance n° 2016-1057 du 3 août 2016 relative à l'expérimentation de véhicules à délégation de conduite sur les voies publiques autorise l'expérimentation de véhicules intelligents sur la voie publique sous condition de la délivrance d'une autorisation accordée par le ministre chargé des transports, après avis du ministre de l'intérieur.</p> <p>Début février 2017, l'Allemagne et la France ont conjointement lancé une zone transfrontalière d'expérimentation pour la voiture autonome entre la Sarre et la Moselle, sur un tronçon Merzig-Sarrebruck-Metz. Le lancement de cette zone d'expérimentation s'accompagne d'une mise en place de guichets uniques pour les demandes d'expérimentations.</p>
<p>Grèce</p>	<p>La Grèce a transposé la directive 2010/40/EU sur le déploiement des transports intelligents, ainsi que le règlement 2015/758 concernant les exigences en matière de réception par type pour le déploiement du système <i>eCall</i> embarqué dans les véhicules. En 2014, en vertu de la décision ministérielle 145/2014, un groupe de travail a été créé afin d'élaborer une stratégie nationale pour les systèmes de transports intelligents. Dans l'état actuel du droit, l'utilisation de véhicules autonomes n'est pas autorisée. Cependant, la Grèce ayant également ratifié la Convention de Vienne, un amendement de la Convention ainsi qu'un amendement des législations nationales qui ont transposé la Convention dans l'ordre juridique interne seront nécessaires pour s'adapter à la réalité de la voiture sans chauffeur.</p>
<p>Israël</p>	<p>La voiture intelligente étant encore au stade embryonnaire, la législation israélienne ne prévoit pas encore de dispositions concernant ces véhicules. Bien au contraire, la section 29 des règlements 5721-1961 sur les transports indiquent qu'un chauffeur doit avoir les mains sur le volant dont il a le contrôle à tout instant.</p>
<p>Italie</p>	<p>Il semble que le législateur n'avait pas envisagé l'hypothèse d'un véhicule sans chauffeur : de fait, le code de la route et des véhicules italien ne pose aucune restriction quant à la manière dont peut circuler une voiture.</p>

Japon	Le code de la route dispose qu'un véhicule doit être conduit par un conducteur maîtrisant et contrôlant sa voiture. De fait, les voitures autonomes ne sont a priori par autorisées à circuler sur la voie publique japonaise. Cependant, il existe un système national de certification des véhicules qui n'exclut pas les véhicules autonomes : si une voiture autonome obtient une certification, elle pourra être expérimentée sur les routes publiques nationales, à condition qu'un conducteur demeure au contrôle du volant, des freins et autres dispositifs de la voiture. Il est également possible d'expérimenter une voiture autonome sur route sans recevoir de certification, grâce à une permission accordée en amont par le ministre.
Liban	L'expérimentation et la circulation de véhicules autonomes ne fait l'objet d'aucune disposition législative ni de texte en écriture.
Portugal	Le Portugal ayant ratifié la Convention de Vienne, les mêmes règles de circulation sont appliquées que dans les pays précédemment cités. Ainsi, le code de la route interdit la circulation de véhicules sans chauffeur, et le chauffeur est responsable de toutes les manœuvres du véhicule. Cependant, de nombreux changements législatifs, au premier lieu duquel l'amendement de la Convention de Vienne, permettront la multiplication des expérimentations et le déploiement des véhicules intelligents.
Royaume-Uni	Le Royaume-Uni ne dispose pas de régulation spécifique concernant les véhicules semi-autonomes ou autonomes - le Royaume-Uni a signé mais n'a pas ratifié la Convention de Vienne. Une révision du cadre juridique national est à l'étude et devrait être présentée au cours de l'année 2017.
Suisse	Le 24 avril 2015, l'entreprise Swisscom, appuyée par le Département fédéral de l'environnement, des transports, de l'énergie et de la communication, a obtenu l'autorisation exceptionnelle de tester un véhicule appartenant à la Freie Universität de Berlin sur des portions de voie publique déterminées. La Suisse n'ayant pas de constructeur automobile national, les données des conducteurs de véhicules autonomes se verront protégés en vertu de la législation européenne en vigueur, notamment la directive européenne pour le déploiement des STI du 7 juillet 2010 ou le règlement général sur la protection des données du 27 avril 2016.

ANNEXE 4 : CONTRIBUTION DE M. RAJA CHATILA, DIRECTEUR DE L'INSTITUT DES SYSTÈMES INTELLIGENTS ET DE ROBOTIQUE (ISIR) « ROBOTIQUE ET INTELLIGENCE ARTIFICIELLE »

La problématique de l'intelligence artificielle telle que posée par Alan Turing¹ était de savoir si les ordinateurs pouvaient être capables de « pensée » (*Can Machines Think?*) et il l'a immédiatement traduite par la question de l'imitation de l'homme. La question initiale de Turing portant sur la « pensée » a été traduite par les fondateurs de l'intelligence artificielle par celle de l'« intelligence », ou des « mécanismes de haut niveau ».

Cette manière de poser la question néglige un constat pourtant simple : le cerveau humain a évolué vers ce qu'il est en développant des capacités de perception, d'interprétation, d'apprentissage et de communication en vue d'une action plus efficace.

La problématique de la robotique pose l'ensemble de ces questions. Le robot-machine est soumis à la complexité du monde réel dans lequel il évolue et dont il doit respecter sa dynamique. La notion d'intelligence doit alors être posée de manière à rendre compte globalement des processus sensori-moteurs, perceptuels et décisionnels permettant l'interaction en temps-réel avec le monde en tenant compte des contraintes d'incomplétude et d'incertitude de perception ou d'action. C'est le sens de la définition de la robotique donnée par Mike Brady (Oxford) dans les années 1980 : « *la robotique est le lien intelligent entre la perception et l'action* ». Dans ce sens on peut aussi dire que le robot est le paradigme de l'intelligence artificielle « incorporée », c'est-à-dire une intelligence matérialisée dans un environnement qu'elle découvre et dans lequel elle agit.

Il est nécessaire d'adopter une vision d'ensemble du robot, en tant que **système** intégrant ses différentes capacités (perception/interprétation, mouvement/action, raisonnement/planification, apprentissage, interaction) et permettant à la fois la réactivité et la prise de décision sur le long terme. Ces fonctions doivent être intégrées de manière cohérente dans une architecture de contrôle globale (architecture cognitive) ; leur étude de manière séparée risque d'aboutir à des solutions inappropriées.

Perception, action, apprentissage

De nombreuses avancées ont été réalisées dans chacune des fonctions fondamentales du robot. Dans les années 1985-2000, la problématique de la localisation et de la cartographie simultanées

¹ A. M. Turing (1950) *Computing Machinery and Intelligence. Mind* 49: 433-460.

(*Simultaneous Localisation And Mapping*) a connu un développement formidable qui a permis de bien en cerner les fondements et de produire des systèmes efficaces, le point faible important restant le manque d'interprétations plus sémantiques de l'environnement et des objets qui le composent. L'apprentissage profond a récemment fait une incursion considérable en robotique, prenant la place de l'apprentissage bayésien, à la fois pour la perception et pour la synthèse d'actions. Mais la perception en robotique nécessite une *interaction* du robot avec son environnement et non une simple *observation* de celui-ci. L'apprentissage par renforcement est un apprentissage non supervisé qui permet au robot de découvrir à la fois les effets de ses actions, caractérisés par une « récompense » obtenue comme conséquence de l'action, et l'incertitude de ses actions qui n'ont pas toujours les mêmes effets. Le lien entre perception et apprentissage - en particulier apprentissage par renforcement - est essentiel pour extraire la notion d'*affordance* qui rend compte des propriétés des objets en ce qu'elles représentent pour l'agent, et qui associe les représentations perceptuelles aux capacités d'action. C'est cela qui sert de base au robot pour exprimer le sens du monde qui l'entoure.

Interaction

Au-delà du traitement du langage naturel, les problèmes d'interaction et d'action conjointe homme-robot, avec la mise en œuvre de « prise de perspective » est une problématique fondamentale pour permettre une interaction efficace et naturelle entre l'humain et le robot. Ce sujet demande un développement qui associe des recherches en robotique et en Sciences de l'Homme et de la Société (SHS), - en particulier, en sociologie, philosophie, psychologie, linguistique. Le rôle des émotions dans l'interaction est à explorer, bien au-delà de travaux actuels qui se contentent de classer *a priori* des expressions faciales ou de produire des expressions d'émotions artificielles par le robot. L'expression d'émotions par un robot pose des questionnements scientifiques et éthiques sur l'authenticité de ces émotions et sur l'anthropomorphisation qui peut en résulter.

Décision

La prise de décision s'appuie sur des techniques relativement classiques de l'intelligence artificielle. La non-adéquation de la planification déterministe avec les contraintes du monde réel a amené le développement de méthodes basées sur les processus décisionnels dans l'incertain, comme les processus markoviens. Ces approches trouvent cependant leurs limites dans une trop grande complexité. L'apprentissage d'actions en robotique est souvent réalisé par des méthodes d'apprentissage par renforcement et le lien entre planification et apprentissage n'est pas très clair. Des recherches sur de nouvelles approches alliant apprentissage et raisonnement probabiliste avec une convergence rapide sont nécessaires.

Conscience de soi

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

Pluridisciplinarité

La recherche en robotique pose des questions proches des sciences cognitives, des neurosciences et de plusieurs domaines des Sciences de l'Homme et de la Société (SHS), comme la sociologie, la psychologie et la philosophie. Des programmes interdisciplinaires seront probablement le bon moyen pour aborder les différentes facettes des questions fondamentales.

Deux aspects concernant l'éthique devraient être abordés: d'une part les *méthodologies de conception* éthique de systèmes autonomes, de manière à ce que ceux-ci tiennent compte des valeurs éthiques humaines (par exemple respect de la vie humaine, des droits humains) et de manière à ce que les algorithmes qui les régissent soient transparents, explicables, traçables, et d'autre part l'éthique des machines, c'est à dire comment les décisions prises par une machine peuvent intégrer un raisonnement éthique.

ANNEXE 5 : CONTRIBUTION DE M. PATRICK ALBERT, ENTREPRENEUR ET CHERCHEUR : RELEVER DÉMOCRATIQUEMENT LES DÉFIS DE L'INTELLIGENCE ARTIFICIELLE

Le développement de l'intelligence artificielle, favorisé par une concentration industrielle démarrée à la fin des années 2000, **bouleverse nos modes de vie, que ce soit au niveau du travail, de l'enseignement, de la santé ou des relations humaines**. Ce bouleversement est tel que chacun, glissant de manière continue vers une dépendance de plus en plus importante aux technologies, produits et services numériques majoritairement importés, peut expérimenter les bénéfices des services de l'intelligence artificielle. Ce développement a connu **trois vagues industrielles successives**, entrecoupées par deux « hivers » de l'intelligence artificielle » faits de désillusions. La vague des années 1980 correspond à l'apparition des « systèmes experts », des applications automatisant la connaissance d'un expert spécialisé dans la réalisation d'une tâche particulière. Cette première vague a vu la création de nombreuses jeunes entreprises dont les services ont permis le développement d'applications intelligentes de diagnostic de pannes ou de contrôle de procédés industriels, par exemple. Cependant, cette avancée technologique ne fut pas assez significative pour permettre à l'intelligence artificielle de connaître une percée industrielle durable. Cette vague fut également marquée par l'utilisation des machines américaines Lisp, du nom du langage informatique standard de développement des algorithmes, et des machines japonaises automatisant la mise en œuvre du langage PROLOG – pour PROGRAMMation LOGique – inventé en France par le professeur Alain Colmerauer. Ces machines n'ont pas cependant dépassé le stade de super-prototype, et sont restés trop onéreuses pour déboucher sur des applications industrielles massives. Une autre vague débuta dans les années 2000, après plus d'une décennie « d'hiver ». Au cours de cette période, les technologies issues des systèmes experts ont connu des progrès considérables, améliorant leur performance et surtout simplifiant à l'extrême leur mise en œuvre par des non-informaticiens. La technologie s'est focalisée sur la simplification extrême de l'automatisation de *décisions simples*, telles que la décision d'allocation d'un prêt, l'application d'un programme de fidélisation client, ou encore l'évaluation de la conformité à une norme. Cette deuxième vague de l'intelligence artificielle est celle de l'automatisation des règles métiers. Ces systèmes permettent au spécialiste d'un métier de transférer à la machine les règles que lui-même met en œuvre en programmant des applications de décisions. Ce procédé se répandit dans de nombreux secteurs, en tout premier lieu dans celui des banques et des assurances.

Les **données** sont généralement considérées comme l'**or noir** d'Internet ; c'est également le cas de l'intelligence artificielle, qui se nourrit des milliards de données fournies en quantité croissante et à un coût marginal. C'est dans ce contexte que la dernière vague de l'IA, dans la période actuelle, émerge, grâce à l'articulation des progrès spectaculaires de la recherche fondamentale ayant perfectionné les premières méthodes d'apprentissage avec la massification des données et l'apparition de calculateurs spécialisés à très bas coût, les processeurs graphiques (GPU)¹. Ces puces GPU rencontrent une demande forte, près de 500 millions d'unités ayant été vendues au cours de l'année 2014 ; elles servent à appliquer des traitements très simples mais simultanément à une grande quantité de données. Intel a ainsi perdu sa position dominante en faveur de nouveaux fabricants de microprocesseurs, tel que Nvidia.

À l'instar de l'informatique qui lui sert de moteur, l'intelligence artificielle est une **technologie horizontale**, elle peut servir « à tout ». Elle fournit aux applications informatiques la capacité de résoudre des problèmes complexes. Elle permet à des machines de **percevoir** et de **décider** pour automatiser des tâches jusqu'ici entièrement réalisées par des humains, ou pour effectuer des activités nouvelles – par exemple les jeux vidéo, ou comme dans le cas de la réalité virtuelle, inventer des mondes dignes de la science-fiction et abuser nos sens pour nous faire croire à leur réalité.

Les intelligences artificielles sont le plus souvent sans mémoire, ou plus précisément, sans autre mémoire que celle – limitée mais parfaite – qui est indispensable pour mener à bien la tâche à laquelle elles sont dédiées. Car elles fonctionnent encore en silos, chacune d'entre elles automatise une tâche très spécifique. Une IA est totalement incompétente en dehors de « sa » tâche et impossible à connecter utilement à des IA développées, pour d'autres besoins dans d'autres contextes. Il existe des langages généralistes de communication inter-agents, mais ils en sont restés au stade de l'expérience de laboratoire. Un langage universel pour les agents du *web* se développe toutefois depuis l'année 2007. Ce qu'on appelle le « *web sémantique* » est une initiative d'envergure mondiale portée par Tim Berners-Lee, l'inventeur du *web*, qui vise à donner aux agents intelligents la capacité de représenter le monde et de le comprendre ; de lui donner du sens. Les langages RDF et OWL² sont ainsi la synthèse de deux décennies de recherche sur les langages de représentation des connaissances, ils fournissent aux agents du *web* un moyen de communication universel indépendant des langages humains – une sorte d'Esperanto³ pour agents intelligents.

¹ En anglais, Graphics Processing Unit. Ce sont des accélérateurs graphiques conçus pour afficher les images ou les films sur tout type d'écran.

² Ontology Web language. L'ontologie, introduite par le Grec Aristote veut décrire le langage pour décrire le monde sous la forme d'un ensemble de concepts reliés entre eux.

³ Si l'Esperanto est en situation de demi-échec, le succès bien réel de la Langue Internationale des Signes démontre la possibilité d'une langue universelle.

On peut séparer **trois grandes classes d'activités « intelligentes »** : la *perception et l'expression* – voir, entendre, parler, etc. – les *décisions dites de routine*, où l'on applique une connaissance simple, une réglementation, une « politique » prédéfinie, et les *décisions d'optimisation* où l'on cherche à obtenir la meilleure solution à un problème de gestion de ressources limitées et contraintes.

Les décisions d'optimisation sont automatisées grâce à leur formalisation dans des modèles mathématiques¹ par des spécialistes ou par des programmes développés par des experts. Après le secteur de **la défense**, ce sont trois industries par natures globalisées qui ont embrayé le pas : les **transports, les télécom et la finance**. En effet, la **logistique**, le **transport** aérien, routier ou ferroviaire ne fonctionnent que grâce à ces agents artificiels qui sélectionnent les transporteurs, calculent leur trajet optimal, décident du chargement des soutes, du camion ou des wagons et planifient la séquence d'actions à mettre en œuvre pour le réaliser. Plus encore, ils sélectionnent le conducteur ou le pilote, ou encore l'ensemble de l'équipage d'un avion, en respectant les conditions de sécurités, les lois et réglementation et les accords syndicaux tout en optimisant le coût financier ou social. Uber et ses algorithmes, ou encore la voiture robot reléguant le chauffeur au rôle de passager, ne sont que la partie émergée exposée au grand public d'un iceberg dont la profondeur ne cesse de s'accroître. Bien entendu les algorithmes d'optimisation calculent le prix optimal² et le font varier au cours du temps, prenant éventuellement en compte les visites successives de l'internaute repéré sur le site de la compagnie.

L'automatisation des services du *web* – disponibles partout et à tout instant – est une nécessité absolue: quelle est la bonne page à référencer pour Google ou Yahoo ? Que contiennent les photos des murs Facebook ? Ou encore quelles sont les compétences et les souhaits de carrière des personnes connectées au réseau professionnel LinkedIn ? Est-on susceptible de recevoir une allocation ? Quelle est la réduction – ou le « cadeau » – à proposer pour déclencher le passage à l'acte de l'acheteur qui hésite ? Peut-on allouer ce crédit et à quel taux ? Quel accessoire est disponible sur cette voiture, à quel prix et avec quels délais de livraison ? La réponse doit être calculée en une fraction de seconde et présentée *immédiatement* ; seule une machine peut le faire. La logique qui pousse les grands acteurs du *web* vers l'intelligence artificielle, dans **une compétition féroce pour la construction de monopoles**

¹ L'informatique est cousine des mathématiques, et nombre d'informaticiens sont des mathématiciens contrariés. Il n'est pas étonnant que tout comme pour la logique, les mathématiques soient à leur tour mobilisées pour résoudre des problèmes de décision ; les deux approches sont parfaitement complémentaires, les nombres, quasiment divinisés par les Pythagoriciens, ont été chez les philosophes les compagnons de route de la logique. Au fronton de l'Académie de Platon était ainsi écrit « Nul n'entre ici s'il n'est géomètre ».

² Le Yield management qui fixe le prix d'un service en fonction du moment où il est délivré : compagnies d'aviations, hôtels ou encore SNCF.

mondiaux, intensifie et accélère ce mouvement. Les plus rapides gagneront la compétition. Le message est bien résumé par Andrew Ng, quittant Google en 2014 pour devenir directeur scientifique de Baidu, le géant de l'Internet Chinois « *Qui gagnera l'intelligence artificielle gagnera l'Internet, en Chine et autour de la planète !* »¹.

L'*Uberisation* caractérise une évolution majeure de la fourniture de services. Des plateformes « biface » portées par le *web* et également accessibles au moyen de smartphones mettent en contact des clients avec des fournisseurs et gèrent le contrat de service qu'ils peuvent passer entre eux. Elles conçoivent et formalisent un service dont elles garantissent la qualité et prennent à leur charge le paiement ainsi que la gestion des inévitables conflits entre fournisseurs et clients. La première génération de plateformes apparue au début des années 2000 est celle des « places de marché en ligne », qui ne fournissent que la mise en relation avec généralement une ébauche de mesure de « confiance ». Ces places de marché se sont généralisées à de très nombreux domaines et avec trois variantes en fonction du type de populations connectées : *B2B* pour les services inter-entreprises, *B2C* pour connecter des professionnels aux particuliers ou *C2C* pour la mise en contact directe de particuliers entre eux. Le plus bel exemple de ces dernières en France étant « Le bon coin » qui fluidifie la circulation quotidienne d'objets et de services entre plus de deux millions six-cent mille particuliers. La simplification à l'extrême de l'accès au service grâce aux efforts combinés sur les plans du marketing et du design, ainsi que la prise en charge des aspects juridiques et du paiement ont donné naissance à la nouvelle génération des plateformes et à leur succès fulgurant. Blablacar en France en est le meilleur exemple. Selon les services, la **garantie de confiance** peut prendre différentes formes, mais on la retrouve le plus souvent incarnée sous la forme d'un bouquet de garanties telles que le contrat prédéfini, le paiement sécurisé, la sécurité du service, les assurances et la gestion des différends, et bien sûr la « réputation » du fournisseur.

Bénéficiant des dernières avancées technologiques, le service fourni est généralement beaucoup plus simple, de meilleure qualité et moins cher. On a donc expliqué que le *web* est un outil de *désintermédiation*. C'est une fausse approximation : les plateformes sont des outils d'*automédiation*, elles automatisent l'intermédiation qui est loin d'avoir disparue. Cette intermédiation, qui met en contact l'offre et la demande n'a pas été dissoute par quelque magie du *web*, elle a été simplement automatisée, ce sont des algorithmes d'intelligence artificielle animés par des machines qui font le lien. Les sociétés telles que Uber ou AirBnB pour ne citer que les plus emblématiques de ce mouvement de fond, sont les nouveaux intermédiaires, les successeurs des courtiers.

¹ « Whoever wins artificial intelligence will win the Internet, in China and around the world ».

Tout comme il y a deux siècles lorsque les premières manufactures industrialisaient la **production** – en l'occurrence des draps –, les premières grandes plateformes du *web* inaugurent le mouvement de l'industrialisation du **service** et du **courtage**. Ni AirBnB, qui revendique plus de 2 millions de logements dans 34 000 villes à travers le monde, ni Booking.com ne possèdent de logements, et Uber ne possède pas plus de taxis que Blablacar n'a de voiture. Ces plateformes automatisent la mise en relation de l'offre et de la demande. Elles fournissent les garanties suffisantes pour que des personnes se fassent suffisamment confiance pour entrer dans une relation contractuelle, à distance et sans se connaître. C'est ce « *magic-mix* » qui fait le succès mondial et fulgurant des plateformes. Et lorsque l'automatisation du service lui-même est possible, nous verrons très probablement les plateformes le fournir directement, sans autre intermédiaire. Ainsi la trajectoire d'Uber est de remplacer une partie de ses conducteurs par des voitures autonomes¹. C'est un fait encore peu perçu, mais c'est bien **l'intelligence artificielle qui est le moteur de cette industrialisation du service que le web et les mobiles ont rendue possible**. Les conducteurs des voitures de tourisme avec chauffeur (VTC) mobilisés par Uber ne sont pas des professionnels du taxi, ils ne connaissent pas le détail des villes au sein desquelles ils transportent leurs clients. Le conducteur Uber est littéralement « piloté » par l'intelligence artificielle qui calcule le « meilleur » trajet². Cette automatisation du calcul du trajet, qui est le cœur du métier, garantit le meilleur service et permet à des personnes non qualifiées de devenir chauffeurs de taxi ; elles n'ont qu'à suivre les ordres de la machine. On peut poursuivre l'exemple avec la fixation du prix de la course qui est défini dynamiquement en fonction du contexte. Cette méthode de fixation du prix est un sujet classique de l'intelligence artificielle, celui du *yield management* dont le but est de maximiser le chiffre d'affaire ou le profit tiré d'une capacité par un pilotage fin de son prix de vente. En changeant radicalement de domaine, on pourrait facilement faire la même démonstration pour l'uberisation naissante des métiers du droit par les plateformes naissantes des « *Legaltech*³ ». Ces plateformes proposent de mettre en contact un client et son avocat, si nécessaire, tout en automatisant à marche forcée le traitement des cas les plus simples ou les plus fréquents. Les techniques seront différentes, par exemple la compréhension du langage naturel ou l'utilisation de robots conversationnels, mais à terme l'effet sera certainement le même.

L'intelligence artificielle se déploie grâce à l'exécution de logiciels spécialisés sur des machines. Ce déploiement se produit le plus souvent dans

¹ En commençant le 25/10/16 avec la livraison de 50 000 bières par un [camion autonome](#) équipé par la start-up Otto.

² Un problème classique d'IA connu sous le nom barbare de « TDTSP » : Time Dependent Travelling Sales Person.

³ Les start-up qui appliquent agressivement l'automatisation algorithmique aux problèmes de législations.

le *cloud*, sur nos machines de bureau ou à la maison, et de façon croissante sur les mobiles¹.



Mais c'est sans conteste « silencieusement » que l'IA se déploie le plus, hors de notre perception sensible.

Les **industries** soumises à une compétition internationale largement dérégulée, donc très aigüe, ont dû lutter pour diminuer leurs coûts et améliorer les décisions. *In fine* elles ont dû confier à des algorithmes intelligents le calcul des solutions optimales à nombre de leurs problèmes opérationnels. Les industries aéronautiques – compagnies d'aviation et aéroports – et la téléphonie fixe ou mobile ont été pionniers dans la mise en œuvre des technologies de l'intelligence artificielle et de sa cousine jusqu'ici plus discrète, la recherche opérationnelle². Toutefois, de nombreuses industries sont maintenant équipées de logiciels intelligents et ce nombre ne fait que croître. L'éditeur de logiciels ILOG, fleuron de l'IA française, vendu à IBM fin 2009, a revendiqué près de trois mille clients : 2 500 entreprises et plus de 465 éditeurs de logiciels. Un impact énorme sur de très nombreuses applications.

Outre le commerce en ligne, déjà évoqué plus haut, il faut bien sûr citer les **banques** et la **finance**. Le cadre féroce de leurs opérations, mondial et dérégulé, et les objets manipulés – essentiellement immatériels – se prêtent à la perfection au monde de l'automatisation. Le cas particulier du « *Trading Haute Fréquence* » est plus que parlant : seuls des robots peuvent entrer dans la compétition car la vitesse de décision est telle qu'aucun humain ne peut combattre. Le *trading* algorithmique représente maintenant près de la moitié des transactions financières mondiales, le *golden boy* contemporain est en pur silicone.

Les **assurances** ont (un peu) tardé à s'y mettre car la compétition n'a pas la nature aigüe de la finance. Mais surfant sur la capacité prédictive apportée par les *big data* et par la connaissance individuelle fournie par les capteurs toujours plus nombreux qui nous environnent, les assurances se mettent à l'IA et la course à l'automatisation n'en est qu'à son début. L'intelligence artificielle sera le moteur d'une hyper-individualisation de la

¹ Sept milliards fin 2014.

² Une annonce de la SNCF qui cherche un « Chef de projet en Intelligence Artificielle » sur le site de la recherche opérationnelle française indique que ces frontières traditionnelles s'émeussent.

gestion du risque assurantiel rendue possible par la combinaison d'une catégorisation des comportements individuels, issue de l'analyse temps-réel des données du *web* et de l'Internet des Objets, avec les modèles prédictifs du *big data*.

Toujours dans le domaine de l'immatériel, les **réseaux sociaux** tels que Facebook utilisent l'intelligence artificielle pour étiqueter les photos publiées¹ ou encore contrôler et définir le flux d'information et l'ordre dans lequel il est présenté. De fait Facebook devient un éditeur qui sélectionne et ordonne les informations soumises à ses clients. Son algorithme fait l'office du rédacteur en chef d'un quotidien unique à chaque lecteur dont il aurait une connaissance intime. De son côté, la **presse** traditionnelle passe en ligne, et mobilise l'intelligence artificielle pour sélectionner des informations pertinentes et rédiger les articles les plus simples. La vitesse de réaction/rédaction est encore ici un atout essentiel.

L'**agriculture** dite « de précision » qui est alimentée par les données GPS et satellitaires² est elle aussi pilotée par intelligence artificielle. Planification et ordonnancement des travaux à réaliser dans les champs, réalisation de plus en plus automatisée de ces travaux, et jusqu'au choix des qualités et quantités d'engrais à déposer sur chaque m² d'un champ par analyse des photos satellitaires.

En s'automatisant la **médecine** produit des traitements intégrant les progrès les plus récents de la connaissance médicale et adaptés au plus près de chaque patient. De son côté, la **pharmacie** peut vérifier la compatibilité des médicaments prescrits tout en s'assurant de la validité des doses. La médecine devient préventive en confrontant les données issues de nos capteurs à des modèles prédictifs découverts grâce à l'analyse des masses de données récoltées sur les cohortes, tout en diminuant le coût d'interprétation des prélèvements et examens dont l'interprétation est partiellement ou totalement automatisée. Les **assurances** vont aussi être des acteurs majeurs de la poussée vers plus de prévention et de personnalisation.

L'automatisation du **marketing** et la **publicité**, on l'a vu, rapprochent le client et le vendeur, en personnalisant l'interaction, diminuant la quantité de publicité hors sujet à laquelle nous sommes exposés et rapprochant l'offre de nos besoins perçus. C'est sur ce domaine d'application qu'est né le *big data* (si on enlève les logiciels réalisés par les géants américains pour les besoins de la NSA). La quantité de publicité contenue dans les pages *web* a vite posé le problème de sa pertinence. Les internautes ont massivement rejeté le tsunami de la e-pub, et la lutte est maintenant ouverte entre les sites « gratuits » financés par la publicité et les

¹ Entre 1 et 1.5 milliard par jour... impossible à faire manuellement.

² Vérifié personnellement par une séance dans le tracteur de l'agriculteur de mon petit village d'Episy dans le Gâtinais. Son intelligence artificielle trace des sillons plus précis que ceux du céréalier.

ad-blockers qui suppriment les publicités. La tentative de compromis en cours entre le public et les publicitaires s'articule autour d'un concept de smart-pub ; une publicité intéressante et pertinente. Les algorithmes d'apprentissage que nous nourrissons de nos traces sur l'Internet servent à caractériser l'internaute et sont le garant de la pertinence. L'enjeu est de taille car c'est le modèle du *web* « gratuit »¹ qui est mobilisé.

Le monde des loisirs n'est pas en reste, en particulier l'univers des **jeux** qui a été l'un des premiers, avec les **militaires**, à investir sérieusement le potentiel de l'intelligence artificielle. Les jeux à deux joueurs se sont très tôt avérés prêter un terrain d'expérimentation presque idéal aux techniques de l'IA. Leur univers extrêmement simple régi par un petit corpus de règles et doté d'une mesure de succès objective leur ont permis de servir de *benchmark* tout autant que de canal de communication avec le grand public. On sait l'écho qu'ont successivement rencontré la victoire aux **échecs** de Deep Blue contre Kasparov en 1997, la victoire au Jeu **Jeopardy**² par IBM en 2011, à l'occasion de son 100^e anniversaire, puis enfin , en mars 2016, la victoire très largement commentée au jeu de **Go** d'AlphaGo, l'IA de la start-up anglaise achetée par Google, DeepMind. Les **jeux vidéo** ne sont pas en reste et utilisent de l'intelligence artificielle depuis plus d'une décennie en accumulant une expérience unique relative à l'importance critique du réglage de l'autonomie de leurs IA : pas assez d'autonomie et c'est l'ennui, mais trop d'autonomie et le jeu, donc le *Gameplay*³, devient imprédictible.

La **réalité virtuelle** peut être la prochaine industrie grosse consommatrice d'intelligence artificielle, car pour être intéressant un univers doit être doté d'un comportement intelligent donc être animé par des IA.

La **musique** est aussi concernée⁴, on peut citer deux exemples de réussites françaises : l'algorithme de flow de **Deezer** qui enchaîne un flux ininterrompu de musiques « qu'on aime »⁵ ou encore l'excellent **générateur de mélodies** pop⁶ créé au sein du laboratoire parisien de Sony Research.

¹ Bien résumé par la célèbre phrase : « Quand c'est gratuit c'est que vous êtes le produit. »

² Équivalent américain de "Question pour un champion ».

³ Ce qui fait que le jeu reste intéressant, voire addictif.

⁴ En matière de musique, il devient difficile de distinguer un morceau fait par l'humain d'un autre entièrement composé par une machine. Dès les débuts de l'intelligence artificielle la composition musicale était dans la ligne de mire. En particulier les années soixante-dix ont vu la montée en puissance des synthétiseurs faciles à programmer et de la notion MIDI propice à son informatisation. Ray Kurzweil, pour n'en citer qu'un, fut un des pionniers de l'application de l'IA à la composition. Il a conçu dans les années soixante des programmes de composition lui valant la célébrité et de nombreux prix. Plus récemment, en décembre 2016, on a pu écouter le morceau « Daddy's Car » lors d'une grande soirée à la Gaité Lyrique, célébrant les vingt ans du laboratoire IA parisien de Sony. Cette excellente chanson qu'on pourrait dire « des Beatles » si cela n'était qu'elle a été composée par la Flow-Machine, une IA conçue par les chercheurs de Sony et éduquée par apprentissage. Ayant attentivement « écouté » et analysé quarante-cinq chansons des Beatles, l'algorithme a appris le style Beatles.

⁵ Notre usage quotidien de cette IA nous laisse parfois encore impressionné par la qualité des propositions de l'algorithme.

⁶ On peut écouter par exemple l'excellent morceau néo-Beatles [Daddy's car](#) sur le site du CSL

La **peinture** n'échappe pas à cette logique : on connaissait déjà les effets graphiques utilisés pour pimenter ou simplement améliorer les millions de « *selfies* » des utilisateurs de Facebook. On peut maintenant faire réinterpréter ses photos « *dans le style de* ». La jeune pousse russe Prisma propose par exemple d'appliquer le style Van Gogh ou Picasso à de simples photos¹. Plus intéressant, en avril 2016 la chaîne de télévision Arte a créé en direct un « nouveau Rembrandt² » qui a véritablement épaté, entre autres, une équipe de professionnelles expertes de la restauration de tableaux anciens. Il ne s'agit plus seulement d'appliquer un style à une composition picturale déjà faite, produite par un appareil photo, mais de créer véritablement une œuvre originale. Avec son projet Magenta³, la firme Google, toujours à la recherche de support médiatique s'est emparée milieu 2016 du sujet de la création graphique. Un premier projet d'art visuel avait suscité un certain écho. DeepDream⁴ produit des représentations graphiques dans des styles si bizarres qu'on pourrait les croire sorties d'un rêve⁵.

Les **robots** connaissent aussi un essor spectaculaire et leur aspect frappe l'imagination : notre tendance à « humaniser » les animaux ou les objets nous incite à intégrer les robots comme une nouvelle espèce « bizarre » qui se développe sous nos yeux produisant parfois un mélange ambigu d'attraction et de peur. Les robots sociaux tels Nao ou Pepper de la société Aldebaran sont conçus pour susciter de l'empathie, et les résultats les plus avancés des neurosciences et de la sociologie sont utilisés pour produire des objets « attachants », des compagnons obéissants fournissant par exemple aux personnes âgées des services de proximités.

L'utilisation d'outils et si possible **l'automatisation du travail** est consubstantielle à l'histoire de l'humanité : il y a quelques 25 siècles Aristote écrivait⁶ : « *Si les navettes tissaient toutes seules, si l'archet jouait seul de la cithare, alors les entrepreneurs se passeraient d'ouvriers, et les maîtres, d'esclaves* ». L'histoire de la relation entre travail, humains et robots, est riche de plusieurs évolutions motivées en premier lieu par la production de masse et le taylorisme. Mais le robot universel reste à inventer. Un robot automatise une tâche particulière. Il est bien plus facile d'automatiser une tâche que l'on aura au préalable simplifiée. Ce n'est pas par hasard que le tout premier robot industriel, *Unimate*, a été mis en œuvre en 1961 sur des lignes d'assemblage, celles de Général Motors. Avec son bras articulé d'une tonne

¹ L'application a été bannie de Facebook probablement pour laisser la place libre à un produit interne encore en gestation.

² Voir : <https://www.youtube.com/watch?v=IuygOYZ1Ngo>

³ Pour une présentation en anglais : <https://magenta.tensorflow.org/welcome-to-magenta>

⁴ L'article de Google : <https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>

⁵ Voir la galerie :

https://photos.google.com/share/AF1QipPX0SCI7OzWilt9LnuQliattX4OUCj_8EP65_cTVnBmS1jnYgsGQAieQUc1VQWdgQ?key=aVBxWjhwSzg2RjJWLWRuVFBBZEN1d205bUdEMnhB

⁶ Aristote, *La Politique*, Livre I, chapitre 2, paragraphe 5.

et demi, ce robot transportait des pièces de fonte de 150 kilogrammes pour les souder sur la carrosserie des voitures en cours d'assemblage. Une tâche ingrate et dangereuse à cause du poids des pièces et des fumées toxiques, mais une tâche extrêmement simple. L'émblématique société Amazon montre la voie du futur¹ : dans une première étape les magasiniers ont été asservis à la machine, travaillant dans le cadre d'une méthode de stockage, connue sous le nom de « Chaotic Storage », conçue par et pour des IA, et uniquement compréhensible par les machines qui les pilotent. Puis dans un deuxième temps, l'entreprise pousse l'automatisation un cran plus loin en déployant des robots en nombre. Ce sont plus de 30 000 robots qui œuvrent aux côtés des humains pour poursuivre l'automatisation du processus de livraison. Plus loin encore, la promesse du *digital manufacturing* tel que développée par exemple en France par la société Dassault Systemes² vise à terme à passer quasi automatiquement de la conception d'un nouveau produit à la « production » de l'usine qui le produira.

Les **GAF**A veulent mener la danse et disposent d'atouts extraordinaires. Ils génèrent **l'équivalent du PIB du Danemark**, la 25^e puissance économique mondiale et sont conçus pour la croissance (de l'ordre de 33 %) et la rentabilité (avec un taux d'imposition de l'ordre de 5 %). Seule la Chine à ce jour a su résister et créer une industrie concurrente crédible, mais leur surface va de l'Internet et du *web* jusqu'à la santé, les médias et le divertissement ou encore la distribution et la finance. Ils capturent **plus de deux milliards d'utilisateurs et plus de 50 % du temps en ligne est passé sur leurs plateformes**. Leur visée est **hégémonique**, la mise en place de monopoles mondiaux, et ils sont activement soutenus dans cette démarche par le gouvernement et l'armée des États-Unis.

Derrière un discours de type ONG les extensions de l'Internet que leurs « fondations » mettent en place semblent viser à **étendre leur marché** – on dit le « *reach* » – et à permettre aux pays en voie de développement d'accéder à l'Internet. À tel point que le peuple indien s'en est ému. Il s'est particulièrement offusqué de l'initiative « Internet.org », promue par un consortium essentiellement américain piloté par Facebook, au titre d'une violation de la « *neutralité du net* ».

L'année 2016 aura vu la concrétisation d'un **mouvement profond et décisif de reconfiguration des GAF**A autour de **l'intelligence artificielle**, après que les années 2000 aient été le moment des expériences et de la consolidation des technologies. Parti de deux projets d'IA en 2011, et après une douzaine d'acquisitions de start-up spécialisées sur le domaine, Google

¹ "Amazon Robotics empowers a smarter, faster, more consistent customer experience through automation. Amazon Robotics automates fulfillment center operations using various methods of robotic technology including autonomous mobile robots, sophisticated control software, language perception, power management, computer vision, depth sensing, machine learning, object recognition, and semantic understanding of commands."

² Voir <http://www.3ds.com/products-services/delmia/disciplines/digital-manufacturing/>

évoque, par exemple, fin 2016 plus de 1000 projets d'intelligence artificielle en cours.

En France, les talents sont nombreux, y compris au sein du monde **académique**¹ et l'entrepreneuriat ne s'est jamais aussi bien porté. Ainsi, coté **start-up**, le dynamisme de la FrenchTech s'avère très efficace². Cependant la **communauté scientifique** de l'IA est morcelée en sous-communautés n'atteignant pas la taille critique qui leur permettrait d'être visibles, il existe peu d'écoles ou d'universités proposant des **cursus solides dédiés à l'intelligence artificielle** et, enfin, la rémunération des **jeunes chercheurs** est plus que modeste. Les **start-up** grouillent, mais comme le résume d'une formule choc Kevin Kelly le fondateur de Wired, les 10 000 prochaines start-up mettront de l'IA dans leur « *marketing mix* »³ et il s'agira d'être convainquant. Faire un peu de *big data* ou d'analyse factorielle ne suffira pas, l'engagement doit être total au risque d'être balayé par des entreprises adossées à une R&D solide. Reste enfin à voir si les dirigeants et les investisseurs vendront les start-up aux GAFAs après les premiers succès ou sauront transformer les PME en ETI puis les ETI en groupes mondiaux à même d'entraîner dans leur sillon un large écosystème de chercheurs et de start-up innovantes.

L'Europe est un **formidable atout** qu'il faut utiliser, mais cela passe par une approche informée des difficultés inhérentes à sa complexité mais aussi par la création préalable d'une communauté nationale forte et consciente d'elle-même : on ne peut sous-traiter la souveraineté scientifique nationale aux institutions européennes.

Il n'y a **pas encore de cadre réglementaire ou juridique spécifique à l'intelligence artificielle** et le déploiement de la technologie avance à un rythme peut-être peu compatible avec celui du législateur. Mais il est ou sera nécessaire d'étendre la régulation aux problèmes posés, qui sont nombreux et parfois critiques : déterminer ce qui peut ou ne peut pas se faire, qui endosse les responsabilités et quelles seront les éventuelles sanctions. Au-delà de la simple régulation, on peut se poser la question de l'utile et s'interroger sur ce qui est bon, ce qui est plus difficile et probablement moins consensuel. Y-a-t-il une limite à fixer à l'automatisation du travail ou à la gestion de notre santé et plus généralement de nos existences ? La relation

¹ Les **institutions** en charge de la recherche et de l'innovation son également pour partie en panne ou à la traîne. Du côté des grands instituts, il apparait que seul **Inria** prend vraiment conscience de l'enjeu et compte ses forces. Le **CEA** semble amorcer le début d'une prise de conscience, mais le **CNRS** de son côté ne promeut l'intelligence artificielle en première place dans aucune de ses sections.

² En 2016, les entreprises de la Tech française ont atteint un nouveau record en passant le cap des 1,5 milliard de dollars de financement. C'est ce que met en avant le rapport « [Tech Funding Trends in France](#) » réalisé par La French Tech et CB Insights.

³ "In fact, the business plans of the next 10 000 start-up are easy to forecast: Take X and add AI".

entre l'homme et la machine, la tendance de la technologie à s'auto-engendrer et à englober toujours plus de pans de nos existences ont été jusqu'ici des sujets plutôt marginaux. Ces sujets ont toutefois été pensés et travaillés par un petit nombre de philosophes alertant le plus souvent sur les dangers et les charmes hypnotiques du confort apporté par une automatisation toujours plus poussée. Selon Satya Nadella, P-DG de Microsoft :

- L'intelligence artificielle doit être conçue pour assister l'Humanité. Construire des robots se chargeant des tâches dangereuses protégerait ainsi les travailleurs humains.

- L'IA doit être transparente. Les principes et règles des algorithmes doivent être explicites et compréhensibles par chacun : le design doit fonctionner de pair avec l'éthique.

- Les algorithmes d'IA doivent être auditables pour les modifier au besoin.

- L'IA doit favoriser l'efficacité sans que cela ne se fasse au détriment de la dignité humaine. Ce qui demande l'implication d'équipes de conception reflétant du mieux possible la diversité culturelle et nécessitant la participation de tous. Le secteur technologique ne doit pas imposer les valeurs du futur.

- L'IA doit éviter tout biais pour ne pas favoriser les discriminations.

- L'IA doit être conçue de manière à protéger la vie privée des utilisateurs.

Il faudra également, en tout état de cause, résister au **mythe de la nécessaire objectivité des processus automatisés** lorsqu'on sait qu'il sont conçus par le tandem bien humain et très subjectif de l'informaticien et du marketeur, fermement aiguillonnés dans les coulisses par leurs investisseurs. Il faudra donc **veiller à conserver, et plus probablement à renforcer, le rôle du politique, du débat public et de la délibération démocratique** dans la gestion des affaires technologiques aussi bien qu'humaines.

On ne peut pas traiter de l'intelligence artificielle aux États-Unis – ni ailleurs... – sans faire un bref détour par la *Silicon Valley*. Ce nom est devenu « générique » et un peu magique. C'est comme un label que l'on appose pour garantir la présence d'une certaine culture vouée à l'innovation. On peut bien sûr retrouver cette culture dans d'autres grands centres américains comme à Boston sur la côte Est, autour du M.I.T. et de Harvard ou encore à Pittsburg, autour de l'université Carnegie Mellon, un haut lieu de la robotique intelligente. New York aussi à sa « *Silicon Alley* » qui revendiquait plus de sept milliards de dollars d'investissement en 2015. D'autres pays à leur tour tentent de créer leur « *Valley* », le plus souvent singeant l'original plutôt que d'en capturer et d'en acclimater l'esprit. Mais la *Silicon Valley* reste quelque chose d'unique : le retour en force de « l'esprit

des pionniers ». Faute de nouvelle terre à conquérir, la conquête de l'Ouest par les colons européens a dû s'arrêter au bord du Pacifique, les laissant sur leur faim. Et de même, la folle ruée vers l'or de 1848, une fois les mines vidées et les rivières asséchées a dû finalement s'éteindre en 1856. Cette ruée vers l'or a créé la Californie, décimant les Indiens et attirant vers l'ouest près de trois cent mille personnes issues de tous pays. Elle a marqué durablement **l'imaginaire californien**. L'esprit de conquête, une soif d'aventure et l'espoir d'une fortune rapide dans un pays dont même les lois étaient encore à inventer. Ce nouveau mythe a déplacé le mythe américain inaugural, celui du fermier travaillant dur son modeste lopin de terre, nourrissant sa famille et servant pieusement Dieu et sa communauté. Mais ce nouvel esprit aventurier est resté brûlant, telle une braise sous la cendre, prêt à être remobilisé. Un peu plus d'un siècle plus tard, la Californie repart à l'attaque ; le silicone a cette fois remplacé l'or, et la ruée a recommencé. L'avènement du monde virtuel de l'Internet suivi de celui du *web* a servi de *nouveau* « nouveau monde » pour réactiver l'imaginaire pionnier. Et la *Silicon Valley* attire à nouveau les aventuriers de tous pays, d'Europe bien sûr mais aussi nombreux sont les Indiens et les Chinois¹, et ce nouveau monde de surcroît est beaucoup **plus accueillant**. Vraiment inhabité, personne à tuer pour occuper ou conserver les territoires, et illimité : dans le monde de l'Internet, la **place est infinie**, il suffit d'incrémenter un numéro pour y planter son ranch, et dans celui du *web* il suffit d'en inventer le *nom* pour créer son territoire. Par contre, revers de la médaille, le territoire nouveau est nu, il y fait nuit et il est inconnu de tous. Dans ce nouveau monde, c'est le diplôme du PhD, le doctorat en informatique, ou le master en marketing qui sert de passeport. Et si la capacité à travailler énormément reste nécessaire, l'esprit innovation est maintenant au pays du silicone ce qu'était le tamis des orpailleurs d'antan.

La Silicon Valley dispose enfin d'un atout qui pour être *soft* n'en est pas moins décisif, elle a parfaitement compris les quatre moteurs à la base de la révolution numérique :

1. L'informatisation d'une activité peut totalement la redéfinir, **rebatte les cartes** et *in fine* restructurer l'écosystème industriel qui la portait.
2. Un **informaticien est un artisan**² – tel un chirurgien – pas l'ouvrier anonyme d'une « usine à logiciel » comme on a voulu le faire croire en

¹ A tel point que lors du recul économique de 2002, les termes consacrés B2B – Business to Business – et B2C – Business to Customer – sont devenus « Back to Bombay » et « Back to China »

² La gestion des employés repose sur un modèle post-industriel où une large part des tâches de surveillance et de contrôle réalisées jusqu'ici par le contremaître ou le petit-chef est réduite ou transférée à des machines émotionnellement neutres. Le management peut alors s'estomper et laisser de la place à l'émergence d'un leadership qui n'est plus basé sur l'autorité – le pouvoir conféré par la position dans l'organigramme – mais sur la compétence technique ou émotionnelle. La structure de commandement qui descend du conseil d'administration vers chaque employé reste très solide bien sûr, mais elle se limite à la transmission des consignes, laissant la place à une certaine auto-

France par exemple. Ce sont des développeuses et des développeurs, celles et ceux qui inventent et écrivent le code des algorithmes, et non pas des « programmeurs » comme on le lit souvent.

3. L'enjeu du *web* est la création de **monopoles mondiaux**.
4. On peut **ne pas respecter les lois**. Et si l'on a suffisamment d'argent, de cynisme et d'intimité avec la classe politique, on peut les changer.

À ce titre, l'industrie des plateformes est une illustration criante. Grâce à une maîtrise à la fine pointe de l'informatique et de l'intelligence artificielle elle peut révolutionner des industries aussi stables que l'hôtellerie ou le transport de personnes, les industrialisant *de l'extérieur* et capturant au passage une marge exorbitante.

Dans le cadre de l'utilisation de l'intelligence artificielle pour le bien de l'humanité, on voit donc que **les logiciels libres** - *l'Open Source* - peuvent être un paravent nécessaire, et peut-être efficace, contre la confiscation des algorithmes par une petite minorité de propriétaires. Au-delà des affichages philanthropiques et des proclamations bienveillantes, la réponse des logiciels libres est aussi un très bon moyen d'**attirer les meilleurs développeurs**, qui y trouvent un environnement au sein duquel ils peuvent travailler librement avec leurs pairs, communiquer leurs résultats et faire avancer la science, tout comme au sein d'un laboratoire hébergé par une université. L'approche ouvre également à l'industriel la porte à la mise en place de standards *de fait* favorables au développement de ceux qui en ont pris l'initiative. Enfin les entreprises à la base de ces systèmes bénéficient gratuitement de toutes les idées et améliorations apportées par les développeurs bénévoles qui expérimentent avec les algorithmes ou qui les étendent. C'est ainsi que l'on a vu les deux géants, Google avec « **TensorFlow** » puis Facebook avec « **Torch** », mettre à leur tour une partie de leurs services d'intelligence artificielle en *Open Source*. IBM a suivi, mais plus modestement, en publiant un standard assez peu utilisé, probablement pour ne pas être de reste et tenter d'afficher un support technique à un discours marketing très offensif sur l'informatique dite « cognitive ». Enfin, dernier en date, c'est au tour du géant chinois Baidu en septembre 2016, de mettre à disposition des

organisation quant à leur mise en œuvre. Les outils informatiques modernes que maîtrise parfaitement la Silicon Valley favorisent une collaboration horizontale, entre pairs, et une poussée vers l'auto-organisation des groupes de projets qui s'accompagne bien sûr d'une plus grande motivation, de meilleurs résultats. L'industrie du logiciel s'inspire en fait largement du mode d'organisation qui a fait l'incroyable succès des communautés scientifiques, tout en y ajoutant la promesse de très fortes rémunérations. Ce mode d'organisation attire les « Smart Creatives » (concept introduit par des entreprises telles que Google et Netflix), dont cette industrie a tant besoin et qui font la différence. En dehors de cette sphère « high tech », dans les industries et services « classiques », celui qui fait est bien souvent nettement moins rémunéré et moins considéré que celui qui commande. Les managers raflent la mise. Le salaire des développeurs vedettes peut dépasser le million de dollars annuel, sans parler de la valorisation des stock-options ou des actions gratuites données à toute personne clé. C'est en particulier un symptôme catastrophique français bien connu qui fait qu'un management préhistorique transforme si souvent d'excellents développeurs en chefs incompetents ou tyranniques.

développeurs son système « Paddle-Paddle » qui regroupe ses algorithmes de *deep learning*.

Nous sommes peut-être à l'aube de la quatrième des *Grandes Transformations* qui ont changé le cours de l'humanité. Ces transformations radicales se sont produites lorsqu'une *ressource critique* jusqu'alors rare et coûteuse d'accès est devenue tout à la fois *abondante* et *abordable* :

- **Nourriture** : l'agriculture, inventée il y a dix à quinze mille ans, a sonné le départ de la croissance des populations humaines, de l'écriture et des technologies.

- **Connaissance** : la presse à imprimer, au milieu du quinzième siècle, a créé en Europe les conditions de l'accès à la culture pour tous, donnant naissance à la sortie de l'obscurantisme et à la civilisation de la connaissance qui a pris son envol avec les Lumières.

- **Énergie** : la machine à vapeur a déclenché au dix-neuvième siècle le mouvement de remplacement du monde des paysans et des artisans par celui du commerce et de l'industrie. Cette domestication de l'énergie a rendu possible la production de masse et la société de quasi-abondance dans laquelle se trouvent en particulier les pays occidentaux.

- **Données et IA** : la quatrième Grande Transformation, dont nous assistons à la naissance, est celle de la donnée, qu'elle soit relative au monde, à nos comportements en ligne ou à ceux captés sur nos existences incarnées. L'information qui était rare, chère, secrète et jalousement protégée, devient surabondante et gratuite. C'est un flux qui se libère, d'aucuns l'appellent un tsunami, mais il est incessant et toujours plus vaste. Ce nouvel or noir est le carburant de l'intelligence artificielle, son omniprésence et sa quasi-gratuité signalent probablement l'aube d'une nouvelle ère.

Chacune de ces grandes transitions apportant des progrès considérables a toutefois bouleversé les sociétés qui les ont vues naître les faisant progressivement disparaître pour finalement les remplacer par une nouvelle. Émergeant au sein d'une société particulière, chacune s'est largement imposée aux autres sociétés, parfois naturellement de par les bienfaits qu'elles apportaient, mais le plus souvent **grâce à la force supérieure qu'elle conférait, utilisée pour l'imposer**. Ces transitions se sont donc souvent accompagnées de misères, de troubles de tous ordres, de guerres civiles ou entre États souvent longues.

Afin de **bénéficier des bienfaits apportés par l'intelligence artificielle** et d'en gérer les risques au plus près, je propose le développement de trois axes complémentaires :

- Construire une communauté nationale de recherche publique et d'innovation industrielle en intelligence artificielle et robotique coordonnée, consciente d'elle-même et réellement compétitive à l'échelle mondiale.

- **Inciter et aider nos entreprises et administrations** à mettre en œuvre ces technologies afin de gagner en efficacité et en compétitivité grâce à l'optimisation des coûts et à l'amélioration de la productivité.

- **Anticiper et gérer démocratiquement**, en impliquant largement la société civile, **les conséquences** de l'avènement de l'intelligence artificielle et de la robotique afin d'en **maximiser les bénéfices pour tous et d'en maîtriser les risques**.

La création d'une « **Agence française de l'intelligence artificielle** » serait un premier pas nécessaire, permettant de regrouper au sein d'une structure dense, experte et focalisée, le traitement des sujets vitaux dont certains ont été rapidement évoqués plus haut : le **développement d'un écosystème national organisé et ultra-compétitif**, à même de se déployer au niveau européen et international, la **promotion et la mise en œuvre de l'IA par nos entreprises et administrations** afin de favoriser leur compétitivité, et enfin la **compréhension fine des enjeux sociaux et sociétaux** pour que le pays bénéficie de promesses de l'IA tout en maîtrisant les risques. Très probablement le troisième Programme D'investissement d'Avenir en cours de finalisation pourrait être un véhicule propice pour porter et financer une telle activité, tout comme les nouvelles ambitions du Plan Juncker pourraient l'accompagner au niveau européen.

ANNEXE 6 : RÉPONSES ÉCRITES AUX RAPPORTEURS

I. RÉPONSES D'AXELLE LEMAIRE, SECRÉTAIRE D'ÉTAT CHARGÉE DU NUMÉRIQUE ET DE L'INNOVATION, ET DE THIERRY MANDON, SECRÉTAIRE D'ÉTAT CHARGÉ DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE, AU QUESTIONNAIRE DE VOS RAPPORTEURS

I. Questions communes

1. Y a-t-il une stratégie nationale en matière d'intelligence artificielle ? Le cas échéant, comment le Gouvernement compte-t-il mettre en œuvre cette stratégie ?

- L'intelligence artificielle est considérée comme un sujet de première importance en raison de ses impacts potentiels importants : opportunités économiques offertes par un nouveau marché (émergence de nouveaux acteurs, de nouvelles offres et de nouveaux usages, gains de productivité), transformations des modèles économiques d'entreprises, nouveaux besoins de formations initiales et continues, créations et disparitions de certains types d'emplois, apparition de nouveaux sujets d'ordre éthique (degrés d'autonomie possibles en fonction des usages, niveau de responsabilité des systèmes autonomes...), etc.

- Avec l'accélération des performances des technologies qui sont au cœur du développement de l'intelligence artificielle (puissances de calculs de plus en plus importantes à des coûts toujours plus bas, technologies du *big data* de plus en plus matures, augmentation des capacités de stockage, développement de nouvelles méthodes comme l'apprentissage profond), l'élaboration et la mise en œuvre d'une stratégie nationale en matière d'intelligence artificielle est devenue une nécessité (sous l'effet aussi de la compétition internationale).

- Cela est d'autant plus vrai que nous observons une compétition internationale déjà très vive, avec en leaders les pays en pointe que sont la Chine et les États-Unis. La stratégie nationale que nous sommes en train de construire doit aussi répondre à cet enjeu en nous donnant les moyens collectifs d'imposer une certaine vision sociale, économique, politique de l'intelligence artificielle, sans doute différente de celles proposées par ces pays.

- Une première étape a été de s'approprier ce sujet très vaste, en particulier au moyen d'entretiens avec des experts de différents secteurs et de différentes disciplines, afin d'en discerner les enjeux et d'évaluer leur importance. Cette première étape a confirmé la nécessité d'agir sans tarder afin que l'avènement des technologies d'intelligence artificielle ne cristallise

pas des craintes qui vont freiner l'innovation et les créations d'emplois et d'entreprises de demain, mais soit, au contraire, porteur d'une volonté de progrès technologique et social.

- La seconde étape, en cours aujourd'hui, consiste à approfondir la réflexion sur les enjeux éthiques, juridiques, technologiques, scientifiques, et socio-économiques de l'intelligence artificielle, puis à proposer et à prioriser, au regard de ces enjeux, des recommandations qui doivent permettre de maximiser les effets positifs du développement et de la diffusion de l'intelligence artificielle en France, tout en amortissant autant que possible les impacts négatifs. À cette fin, en octobre 2016, la ministre chargée du Numérique et de l'Innovation a souhaité la mobilisation de toutes les parties prenantes dans l'établissement d'une feuille de route pour l'intelligence artificielle qui devra être finalisée fin mars 2017. L'écosystème de l'enseignement supérieur et de la recherche est bien sûr autant concerné que celui des entreprises et de l'innovation. Axelle Lemaire et Thierry Mandon, secrétaire d'État chargé de l'enseignement supérieur et de la recherche ont ainsi procédé au **lancement officiel du projet # France intelligence artificielle le 20 janvier 2017**. La méthode retenue consiste en un travail coopératif entre acteurs de l'intelligence artificielle (chercheurs, entreprises, économistes, etc.) et de l'administration (DGE, DGRI, DGEIP, CGI, DG Trésor), au sein de groupes de travail rassemblés en trois grandes thématiques clés. Une première étape de restitution des travaux à mi-parcours a eu lieu le 21 février 2017. L'ensemble des livrables attendus des groupes de travail seront remis le 14 mars 2017. La présentation de la feuille de route issue de ces travaux est programmée le 21 mars 2017.

- Une troisième étape consistera à déployer les actions retenues dans la feuille de route, avec l'appui des financements annoncés par le Président de la République le 21 février dernier.

2. Quelles initiatives le Gouvernement compte-t-il prendre pour relever les défis de l'intelligence artificielle ? Pouvez-vous nous les décrire ?

- D'ores et déjà, comme mentionné précédemment, différentes initiatives ont été prises pour relever les défis de l'intelligence artificielle : entretiens avec des acteurs clés du domaine, définition d'objectifs prioritaires (placer la France sur la carte du monde de l'intelligence artificielle, mobiliser la communauté de l'intelligence artificielle en France et la cartographier, mobiliser des financements et construire un soutien complet à l'intelligence artificielle de la recherche amont aux start-up, porter la question de l'intelligence artificielle dans le débat politique et clarifier le discours auprès des citoyens notamment, en prenant en compte les dimensions éthiques, juridiques, technologiques, scientifiques, et d'impact socio-économique), et lancement de l'élaboration d'une feuille de route pour l'intelligence artificielle.

- Depuis plusieurs années, le développement de technologies et d'usages de l'intelligence artificielle fait, d'autre part, l'objet d'un soutien au travers d'appels à projets de R&D qui permettent des coopérations entre laboratoires de recherche, entreprises fournisseurs de technologies et entreprises utilisatrices. Des projets ont ainsi été soutenus dans le cadre du concours d'innovation numérique du programme d'investissements d'avenir.

- Les groupes de travail constitués pour l'élaboration de la feuille de route de l'intelligence artificielle vont proposer des actions et recommandations sur différents sujets, notamment la maîtrise de thèmes de recherche amont, les formations à créer ou à faire évoluer, la question du transfert des résultats de la recherche aux entreprises, le développement de l'écosystème des fournisseurs de technologies de l'intelligence artificielle, ou encore l'accompagnement des transformations sociales induites par l'intelligence artificielle. Ces propositions étant attendues mi-mars 2017, il est aujourd'hui trop tôt pour en donner le contenu précis

3. Comment le Gouvernement se positionne-t-il à la suite de la remise des rapports récents de la présidence des États-Unis et de la Chambre des Communes sur l'intelligence artificielle (octobre 2016) ?

- Les potentiels, enjeux et risques de l'intelligence artificielle présentés dans les rapports récents de la présidence des États-Unis et de la Chambre des Communes correspondent à ceux communément mentionnés par les experts. Il est intéressant de souligner la place importante des enjeux d'ordre moral et social.

- Par exemple :

- concernant les enjeux d'ordre moral : influence de l'intelligence artificielle sur la liberté de choix des êtres humains, bases éthiques sur lesquelles asseoir les décisions d'une intelligence artificielle, prise en compte des valeurs citoyennes de la société dans laquelle l'intelligence artificielle se développe ;

- concernant les enjeux d'ordre social : concentration des richesses et creusement des inégalités par un déséquilibre accru entre la rémunération du capital (technologique) et celle du travail, concentration du pouvoir, de la connaissance et des choix sociétaux dans les mains des possesseurs d'une technologie devenue incontournable ;

- Concernant les pistes de réflexion et propositions d'actions de l'administration Obama, on peut noter le souhait louable de ne pas laisser les technologies d'intelligence artificielle entre les seules mains d'intérêts privés, qui pourraient confisquer le potentiel économique de l'intelligence artificielle et biaiser le comportement des technologies en fonction de leurs intérêts financiers. L'administration Obama suggère pour cela un effort public massif et une diffusion ouverte des résultats de recherche sur l'intelligence

artificielle. Il convient cependant de souligner que les six grands acteurs du numérique nord-américains que sont Google, Amazon, Facebook, Apple, Microsoft et IBM, investissent déjà annuellement en R&D plus de 45 Md\$. L'intégralité de ce montant n'est pas consacrée aux technologies d'intelligence artificielle, mais il illustre la puissance de ces groupes et le fait que les intérêts privés de grands groupes internationaux vont nécessairement influencer le développement de l'intelligence artificielle et, sans doute, rendre difficile l'ouverture des résultats de recherche. Aussi, au-delà de l'effort public massif suggéré par l'administration Obama, un encadrement, sur le long terme, concernant la transparence des processus de décisions et l'usage des technologies d'intelligence artificielle sera nécessaire.

- La démarche conduite par l'administration américaine est, d'autre part, intéressante dans sa nature même : l'État doit s'interroger sur son rôle pour permettre un développement de l'intelligence artificielle profitable à l'ensemble de la société. Différents scénarios sont, en effet, possibles et il appartient à l'État de savoir lequel il veut promouvoir pour prendre les initiatives qui sont adaptées. C'est également dans cette optique que la feuille de route pour l'intelligence artificielle a été lancée.

4. Y a-t-il des opportunités que la France pourra selon vous saisir grâce à l'intelligence artificielle ? Quelles sont-elles ?

- Les opportunités semblent aujourd'hui grandes ouvertes dans beaucoup de secteurs (droit, éducation, finances, agriculture, transports, etc.). La France se distingue d'ores et déjà particulièrement sur l'innovation dans les domaines de l'énergie et de la santé, portée par la robotique et les objets connectés. Il convient à ce titre de rappeler que les opportunités en matière d'intelligence artificielle sont autant économiques que sociales et sociétales.

- La France dispose également de plusieurs atouts qui peuvent lui permettre de tirer profit du développement de l'intelligence artificielle.

- L'excellence de la recherche française dans des domaines scientifiques et technologiques clés pour l'intelligence artificielle (mathématiques, informatique, sciences des données) est reconnue internationalement. Cette excellence attire les grands groupes internationaux qui implantent sur le territoire national des centres de R&D spécialisés. L'installation de centres de R&D du groupe nord-américain Facebook et du groupe japonais Rakuten en sont des illustrations.

- Au-delà de la dimension scientifique et technique, de nombreux chercheurs français en Sciences Humaines et Sociales s'intéressent également aux impacts socio-économiques, juridiques et éthiques de l'usage des technologies en général, de l'intelligence artificielle en particulier. C'est à la fois un facteur essentiel à la construction et au développement d'une stratégie d'intelligence artificielle conforme aux valeurs de la société que

nous défendons, et sans doute un vecteur du rayonnement et de la différenciation de la France en matière d'intelligence artificielle.

- Un environnement favorable à l'émergence et au développement des start-up a été progressivement développé et constitue une réussite indéniable. Il est notamment le résultat de l'existence de laboratoires d'excellence, d'universités et d'école d'ingénieurs de très haut niveau, ainsi que de la volonté de l'État d'encourager l'entrepreneuriat et la création d'entreprises, en particulier dans le domaine du numérique avec la French Tech. Le fonds d'investissement ISAI a ainsi recensé plus de 180 start-up en intelligence artificielle en France. Ces start-up contribuent au rayonnement de la France dans le domaine de l'intelligence artificielle, créent des emplois et vont faire naître des leaders et permettre à des entreprises de trouver des relais de croissance.

- La France compte, par ailleurs, de grandes entreprises de services du numérique, de dimension nationale ou mondiale (Capgemini, Atos, Sopra-Steria, etc.) qui investissent dans les technologies d'intelligence artificielle. Ces entreprises pourront notamment profiter de la croissance prévue du marché de l'intelligence artificielle pour les applications en entreprise (marché qui devrait passer de 200 M\$ en 2015 à plus de 11 Md\$ en 2024 selon le cabinet d'analyse Tractica) et contribueront à l'appropriation de ces technologies par leurs clients, en particulier français.

- Enfin, l'existence de très grands groupes français utilisateurs de technologies, qui sont des références dans leur secteur, constitue également un atout. En intégrant au plus tôt des technologies d'intelligence artificielle, ces très grands groupes, comme Renault, PSA, Airbus, BNP Paribas, vont développer de nouvelles offres innovantes qui dynamisera leur croissance et stimulera le marché français de l'intelligence artificielle.

5. Quels sont les principaux risques des technologies d'intelligence artificielle selon vous ?

- Les technologies d'intelligence artificielle sont communément associées à l'idée de systèmes autonomes, c'est-à-dire capables de prendre leurs propres décisions. En fonction des technologies d'intelligence artificielle, la manière dont la machine va prendre une décision n'est pas toujours formellement comprise et maîtrisée. Deux systèmes d'intelligence artificielle pourraient donc prendre des décisions différentes dans un contexte identique, ce qui présente des risques dans de nombreux cas d'utilisation. La poursuite de travaux de recherche sur les sujets liés à la maîtrise des processus de décision par les technologies d'intelligence artificielle est donc indispensable.

- Un autre facteur important de risques serait une intelligence artificielle n'intégrant pas de paramètres d'ordre éthique dans son processus de prise de décision. La définition de bases d'éthique faisant consensus et les

modalités de leur alimentation au fil des évolutions de l'intelligence artificielle, est un sujet à traiter en amont et qui devra faire l'objet d'une attention permanente. La prise en compte des sujets d'éthique nécessitera également de développer des outils permettant de démontrer que ces bases ont bien été intégrées dans les technologies d'intelligence artificielle utilisées. Ces outils devront s'intégrer dans des processus de contrôle qu'il convient également de définir dès à présent, et de faire évoluer en permanence.

- Le développement des technologies d'intelligence artificielle qui va s'accompagner d'une facilitation de leur usage peut, d'autre part, faire peser des risques sur la sécurité des États, des organisations et des citoyens. Mises entre les mains d'acteurs malveillants, ces technologies peuvent représenter une capacité importante de nuisance. Une maîtrise française des technologies d'intelligence artificielle (et pas uniquement de leur utilisation) répond donc à un enjeu de souveraineté et de sécurité nationale.

- Enfin, on ne peut s'abstenir d'évoquer les risques d'accroissement des inégalités sociales associés à l'avènement de l'intelligence artificielle - bien que cela n'y soit pas spécifique - entre les citoyens capables d'y accéder et d'en tirer tout le bénéfice, et ceux qui en resteront exclus. La question de l'inclusion numérique doit à plus forte raison rester une priorité.

6. En matière de régulation de l'intelligence artificielle, quelles sont les priorités ? Faut-il aller jusqu'à légiférer ? Si oui, cela doit-il se faire au niveau national, européen ou international ?

- Les priorités doivent se porter sur les utilisations de l'intelligence artificielle qui posent directement des risques sur la sécurité, la santé et le bien-être des citoyens. La régulation pourrait notamment s'attacher à prévoir les situations dans lesquelles les réactions d'une intelligence artificielle doivent être préalablement connues ainsi qu'à définir les responsabilités liées à l'utilisation d'une intelligence artificielle et aux conséquences de ses réactions.

- Une éventuelle réglementation ou une régulation de l'intelligence artificielle bénéficierait, du fait de la nature et de l'impact de la technologie, d'une échelle la plus large possible : européenne a minima pour les questions éthiques et sociales, pour éviter les pratiques de forum shopping des entreprises, internationale pour les questions d'armements (robots tueurs). Le Parlement européen s'est d'ailleurs saisi du sujet, par l'intermédiaire de sa commission des affaires juridiques, qui a adopté, sur la base d'un rapport, une résolution appelant la Commission à proposer une directive dédiée aux questions de robotiques et d'intelligence artificielle. Les recommandations du rapport (taxe sur le travail des robots, agence européenne de l'intelligence artificielle) mériteraient néanmoins d'être instruites en profondeur, en raison de leur impact social et économique potentiel important.

- Enfin, compte tenu du manque de recul que nous avons aujourd’hui en réalité, la régulation encadrant l’intelligence artificielle devra certainement prévoir des éléments de souplesse qui permettront à la fois des ajustements en fonction des usages et données réelles, et de ne pas décourager les investisseurs par une régulation *a priori*.

7. Est-il suffisant de fixer un cadre éthique sans légiférer ? Si oui, à quelles conditions ? Quel contenu envisagez-vous pour ce cadre éthique ?

- La réflexion autour du cadre législatif permet de porter les questions d’éthique dans le débat public. Des règles d’éthiques admises universellement n’existant pas, il est important que certains sujets, particulièrement sensibles, fassent l’objet d’un débat public permettant une sensibilisation des citoyens aux questions en jeu, le recueil le plus large possible des avis, leur discussion et une prise de décision démocratique garantissant une certaine solidité des résultats obtenus. La question même de ce qui relève de l’éthique en matière d’intelligence artificielle fait aujourd’hui débat. C’est pourquoi l’éclairage de la recherche - notamment en SHS - en la matière est également fondamental. En tout état de cause et comme indiqué précédemment, des processus et outils de contrôle validant la conformité aux bases d’éthique définies collectivement, est une piste opérationnelle.

II. Questions spécifiques pour Madame Axelle Lemaire, Secrétaire d’État chargée du Numérique et de l’Innovation

1. Comment est assuré le suivi des principaux défis que représente l’intelligence artificielle ?

- Au niveau du cabinet, un conseiller est spécifiquement en charge du suivi de ce sujet et de la cohérence des actions qui sont décidées et mises en œuvre. (Peut-être, si opportun, mettre en avant des contacts réguliers avec certains acteurs ?)

- Au sein de la DGE, administration rattachée au ministère de l’Économie et des Finances, l’intelligence artificielle est une thématique à part entière prise en charge par le Service de l’Économie Numérique et à laquelle des agents sont affectés.

- Comme mentionné précédemment, les travaux sur l’élaboration d’une feuille de route sur l’intelligence artificielle ont été officiellement lancés le 20 janvier. En fonction des résultats de ces travaux, un suivi systématisé de certains sujets sera mis en œuvre par le cabinet en lien avec les administrations concernées.

2. Pouvez-vous présenter la French Tech et préciser la place qu'occupent en son sein les entreprises mettant en œuvre des technologies d'intelligence artificielle ?

- La French Tech est un terme générique pour désigner tous ceux qui travaillent dans ou pour les start-up françaises en France ou à l'étranger. Les entrepreneurs en premier lieu, mais aussi les investisseurs, ingénieurs, designers, développeurs, grands groupes, associations, medias, opérateurs publics, organismes de recherche et universités... qui s'engagent pour la croissance des start-up d'une part et leur rayonnement international d'autre part. Le nombre de start-up en France est estimé entre 7 000 et 10 000.

- La French Tech désigne également une action publique innovante pour soutenir et promouvoir cet écosystème des start-up françaises dans l'objectif de favoriser leur croissance et l'émergence de nouveaux grands leaders de l'économie française, créateurs de valeur et d'emplois.

- Des acteurs de la French Tech, investisseurs et entrepreneurs français, sous la houlette du fonds d'investissement ISAI, se sont fédérés en 2016 au sein de l'initiative « FranceisAI » pour recenser et promouvoir à l'international les *startups* françaises utilisatrices de technologies IA : près de 200 start-up ont ainsi été identifiées (franceisai.com). « FranceisAI » participe activement à l'initiative « #FranceIA »

3. Quelles sont les principales opportunités qui vont s'ouvrir pour nos entreprises grâce aux innovations en intelligence artificielle ?

- Le fonds ISAI recense, en France, plus de 180 start-up dans le domaine de l'intelligence artificielle. Ce chiffre illustre les opportunités de développement de nouvelles activités qu'offre l'intelligence artificielle en France, en particulier du fait d'un potentiel technique et scientifique important dans le domaine (excellence de la recherche et des formations) et des initiatives prises pour favoriser l'entrepreneuriat.

- Le dynamisme de l'entrepreneuriat français dans le domaine de l'intelligence artificielle attire l'attention des investisseurs. Plusieurs levées de fonds significatives ont ainsi eu lieu en 2016, notamment celle de 8,9 M€ de Shift Technology et de 5,6 M€ de Snips. Cette semaine, le fonds d'investissement français Serena Capital a annoncé la création d'une structure dotée de 80 M€ pour prendre des participations dans des start-up de l'intelligence artificielle et du *big data*. Grâce à ces différentes annonces, c'est finalement tout l'écosystème des start-up français qui est mis en visibilité et qui bénéficie d'un renforcement de son accès aux financements.

- L'intelligence artificielle offre aux entreprises plus matures l'opportunité de profiter de relais de croissance en développant des offres innovantes de produits ou de services, en particulier dans des secteurs où la France dispose d'une solide notoriété comme l'ingénierie informatique (nouvelles offres liées à l'automatisation de processus, à l'intégration de

machine learning), les services collectifs ou « utilities » (nouvelles offres liées à la « Ville Intelligente »), ou encore la sécurité (nouvelles offres liées aux applications de reconnaissance d'images).

- L'intelligence artificielle va également permettre à nos entreprises de renforcer leur productivité et leurs performances (automatisation des processus, analyses prédictives sur le comportement des consommateurs, maintenances prédictives, etc.).

4. Quelles initiatives spécifiques votre secrétariat d'État entend-il prendre dans les prochaines semaines en matière d'intelligence artificielle ?

- La secrétaire d'État chargée du Numérique et de l'Innovation et le secrétaire d'État chargé de l'Enseignement supérieur et de la Recherche ont officialisé vendredi 20 janvier le lancement des travaux sur l'établissement d'une feuille de route sur l'intelligence artificielle. Un point d'étape à mi-parcours a été effectué le 21 février dernier. Les groupes de travail rendront leurs premiers livrables concernant l'analyse et les enjeux des thèmes qui leur auront été attribués. La fin de travaux est programmée le 14 mars, date à laquelle les groupes de travail remettront leurs propositions d'actions. Un événement de restitution des résultats des travaux est programmé le 28 mars.

5. Quelles incitations peut-on envisager à plus long terme en faveur des entreprises pour y favoriser la recherche et l'innovation en intelligence artificielle ?

- Dans un premier temps, une prise de conscience du potentiel de l'AI et des changements importants qu'elle peut induire est nécessaire. Les travaux de la feuille de route et la communication qui sera faite sur celle-ci vont y contribuer, notamment en portant le sujet de l'intelligence artificielle dans le débat public.

- Concernant plus spécifiquement le fait de favoriser la recherche et l'innovation en intelligence artificielle en entreprise, plusieurs types de mesures peuvent être envisagés comme la création de formations continues sur les technologies d'intelligence artificielle, la création de centres de ressources partagées en intelligence artificielle pour les entreprises, le lancement d'appels à projets de R&D dédiés, etc. Un groupe de travail constitué dans le cadre de la feuille de route sur l'intelligence artificielle est plus particulièrement chargé d'établir des recommandations sur ce sujet et celui de l'appropriation des résultats de la recherche par les entreprises. Ses conclusions permettront de répondre plus précisément à cette question.

III. - Questions spécifiques pour Thierry Mandon, Secrétaire d'État chargé de l'Enseignement supérieur et de la Recherche

1. En quoi l'intelligence artificielle peut-elle permettre de révolutionner et de moderniser l'éducation et l'enseignement supérieur ?

Comme pour tous les autres domaines d'action de l'État (finances, justice, santé, etc.), l'éducation et l'enseignement supérieur produisent leur lot de données dans l'exercice même de leur action. Ces données, lorsqu'elles sont assemblées, peuvent faire l'objet de traitements déductifs ou prédictifs. Elles doivent alors être mises à la disposition des acteurs, avec les outils appropriés, pour moderniser plus que révolutionner les modes d'apprentissage. Cela concerne par exemple le suivi du travail de l'élève ou de l'étudiant, d'anticiper le décrochement scolaire ou de pointer les parties de cours nécessitant un accompagnement supplémentaire. L'approche de l'intelligence artificielle permettra ici de mettre en place une personnalisation de l'apprentissage, de tenir compte de profils d'apprenant de plus en plus diversifiés et contribuer ainsi à l'amélioration de la réussite.

2. Comment s'articulent la recherche publique et la recherche privée en intelligence artificielle aujourd'hui ?

La recherche privée s'oriente essentiellement vers les traitements statistiques des données et vers les modalités de reconnaissances des images, de la voix, etc. Cela concerne autant les grands acteurs d'Internet et des réseaux sociaux cherchant ainsi à valoriser les données acquises sur leurs plates-formes, que les jeunes entreprises innovantes désireuses de mettre au point des applications logicielles exploitant données et nouveaux objets connectés. Il s'agit là d'une forme de maturation des concepts développés dans les laboratoires académiques durant les années passées.

En revanche tout le reste des aspects pourtant essentiels de l'intelligence artificielle (notamment le traitement des langues, les différentes formes de logique, la représentation des connaissances, etc.), en interaction avec les SHS, constituent bel et bien le socle de recherche fondamentale nécessaire à l'avancée de la discipline et du domaine. La recherche académique continue donc de travailler sur des objets amont, sur les ruptures et sur l'interdisciplinarité.

En France, la coopération public/privé en matière de recherche s'effectue sur le terrain. Il existe ainsi des partenariats vertueux entre acteurs académiques et privés autour de la fabrication de nouveaux produits, de la création de start-up, de l'embauche de jeunes diplômés issus de Master, d'écoles d'ingénieurs ou d'écoles doctorales.

3. Pensez-vous que cette place respective de la recherche publique et de la recherche privée doit évoluer ?

L'un des enjeux majeurs de l'intelligence artificielle est de pouvoir concilier autour de mêmes réalisations, plusieurs approches issues des sous-domaines de l'intelligence artificielle. Il y a donc là une véritable place pour une collaboration efficiente entre recherche publique (socle fondamental) et recherche privée (exploitation de données massives).

Ces complémentarités devront faire l'objet d'une coordination stratégique, résultant de l'analyse en cours menée dans le programme #FranceIA.

Deux remarques toutefois. Les temporalités d'action de ces deux acteurs ne sont pas les mêmes ; la recherche privée aura toujours besoin d'un calendrier de résultats, d'annonces, de mises à jour. Le secteur public peut lui aussi s'emparer d'un volet lié à l'exploitation massive de données, notamment dans les secteurs qui lui sont légitimes comme l'éducation, la santé, le développement régional, sans pour autant abandonner ces terrains à la recherche privée.

4. Quelle place la France occupe-t-elle dans la recherche internationale en intelligence artificielle ? Comment cette place va-t-elle et peut-elle évoluer ?

La France occupe une place très enviée dans la recherche internationale en intelligence artificielle. D'une part, notre dispositif d'enseignement supérieur dans le domaine, notamment au niveau doctoral, est très performant et bon nombre de jeunes chercheurs trouvent des propositions de travail, tant dans le secteur des entreprises que dans le secteur universitaire au niveau international. D'autre part, les laboratoires de recherche français occupent tout le spectre des recherches du domaine de l'intelligence artificielle, peu de pays disposent d'acteurs performants sur l'ensemble des sous-domaines de l'intelligence artificielle (cela est visible dans les comités scientifiques des grandes conférences académiques internationales pour chaque sous-domaine). Enfin, les jeunes entrepreneurs français font preuve de vivacité et d'imagination dans les solutions qu'ils proposent à la mise sur le marché.

Les atouts de la France tiennent dans la capacité des acteurs à tirer profit à la fois de l'approche française en termes de mathématiques et de modélisation, et de l'approche française en terme d'ingénierie et de gestion de projet.

Un autre atout français réside dans la pluralité et la grande qualité de la recherche en sciences humaines et sociales. Sociologues, juristes, psychologues, économistes, ethnologues, linguistes, philosophes, etc. sont nombreux, travaillent dans des laboratoires de premier plan et pour beaucoup d'entre eux sont technophiles. Il y a là un gisement pour stimuler

et alimenter des programmes de recherche interdisciplinaires aptes à orienter le monde anthropotechnique qui se dessine.

L'évolution de cette position de la France dans la recherche internationale en intelligence artificielle doit tenir compte de risques réels comme la fuite de cerveaux à l'étranger, le contrôle des entreprises naissantes par des acteurs extérieurs, etc. Le plan ambitieux #FranceIA doit aussi réfléchir aux conditions qui permettront d'améliorer la visibilité et l'attractivité de la France pour les chercheurs en intelligence artificielle.

5. Quels sont les domaines de recherche prioritaires selon vous ?

D'une part, il est nécessaire que la France conserve et poursuive le travail des recherches menées sur chaque sous-domaine de l'intelligence artificielle. Il est actuellement impossible de prédire si les créations de valeur viendront demain du web sémantique et de ses ontologies, ou de l'analyse de la voix et des émotions, etc. Sur tous ces points et sur l'ensemble des sous-domaines de l'intelligence artificielle, la France a une position en pointe et doit la garder.

D'autre part, l'intelligence artificielle n'est plus l'affaire des seuls informaticiens, et il faut aussi solliciter des travaux interdisciplinaires d'envergure, notamment pour relier les SHS et les STIC. L'intelligence artificielle développe des solutions prédictives ou de recommandation soit sur la base de scénarios déterminés, soit à partir de calculs basés sur des exemples déjà rencontrés. En revanche rien ne couvre les domaines imprévus, les changements de tendances, les indécisions humaines, les retournements d'opinion. Là où il y a rupture technologique, il y a aussi rupture de mentalité. En même temps que les recherches sur les technologies, il faut produire des recherches sur les intentions, les détournements, les envies. Plutôt que de parler d'acceptation des technologies, il faut envisager d'emblée l'appropriation de ces technologies.

Dans le cadre du plan intelligence artificielle, un groupe de travail est précisément consacré à la cartographie des activités de recherche en lien avec l'intelligence artificielle et l'identification des enjeux de la recherche fondamentale. C'est sur cette base que nous pourrions aussi mieux orienter les efforts.

6. Pensez-vous que des progrès majeurs puissent être attendus de l'hybridation entre technologies d'intelligence artificielle ?

Indiscutablement, oui. À court terme il va s'agir de travailler sur des associations à partir de sous-domaines actuels de l'intelligence artificielle : vision et audition, robots et décisions, perception et action, représentation des connaissances et traitement de la langue naturelle, etc. Ces associations vont produire des solutions beaucoup plus performantes, rendues possibles

par les capacités d'acquisition et de traitements de données. Elles vont aussi pointer de nouveaux verrous qui alimenteront les questions de recherche.

7. Pensez-vous que les principaux progrès seront réalisés par la conjugaison entre l'intelligence artificielle et l'intelligence humaine ?

Des progrès seront probablement accomplis par la conjugaison entre l'intelligence artificielle et l'intelligence humaine. Un humain avec des outils adaptés sera toujours plus performant qu'un humain seul ou qu'une machine seule. C'est par exemple le cas en industrie avec le développement de la cobotique. C'est aussi le cas pour tous les métiers du « care » (de la santé, de l'assistance aux personnes) où un diagnostic rendu plus rapidement, plus précisément, devra toujours être accompagné d'une décision, d'une action qui prend en compte l'humain de façon empathique.

II. RÉPONSES À LA CONSULTATION ORGANISÉE PAR LE PARIS MACHINE LEARNING MEETUP EN VUE DU RAPPORT DE L'OPECST

Nicolas Beaume (15 février 2017)

Je suis bio-informaticien spécialisé dans l'utilisation des méthodes d'intelligence artificielle (plus précisément d'apprentissage automatique) en génomique. Vu la quantité toujours croissante de données que nous brassons en génomique et la complexité des problèmes abordés, les méthodes d'apprentissages automatiques sont plus qu'utiles, elles deviennent indispensables.

Je pense que ce champ et celui de l'intelligence artificielle en général va devenir un pivot inévitable de nos sociétés. Inévitable ne signifie pas qu'il faut en faire n'importe quoi et c'est pourquoi je trouve cette initiative particulièrement intéressante. Si mon expertise en la matière peut être utile, je serais ravi d'y participer.

Olivier Guillaume, fondateur/CEO d'O² Quant, société spécialisée dans les dernières techniques en intelligence artificielle (17 janvier 2017)

Tout comme l'électricité a augmenté les capacités physiques de l'homme, l'intelligence artificielle augmente de manière significative nos capacités cognitives.

Cette évolution majeure a déjà commencé, on la retrouve déjà dans bon nombres d'applications utiles de la vie courante : smartphones, santé, sécurité, industrie... et ce n'est que le début. Demain ce sera les voitures autonomes de manière certaine, et bien d'autres évolutions encore aujourd'hui insoupçonnées.

Avec l'augmentation très récente des données informatiques (le *big data*) et des puissances de calculs, nous entrons maintenant dans une autre dimension, ou le champ des possibles explose. Avec ma société O² Quant, nous travaillons sur des projets améliorant de manière significative la productivité ou la condition de l'homme : amélioration de la sécurité, aéronautique, industrie, réduction du gaspillage alimentaire, recherche appliqué en santé avec des résultats extrêmement rapides...

Même si l'intelligence artificielle pourra être utilisé à des fins malveillantes, je suis très confiant dans l'avenir et à ce que ces nouvelles technologies pourront apporter à l'humanité. Notre pays où il fait bon vivre a bien des atouts, notamment en tant qu'un des acteurs mondiaux majeurs dans l'intelligence artificielle : Nous avons de nombreuses et très bonnes écoles d'informatique et de mathématique ce qui nous procure les data scientists les plus réputés, le dispositif CIR aide les industriels à faire émerger des projets, et bien d'autres atouts ...

Certains pays comme les États-Unis, le Canada ou l'Allemagne croient et investissent dans cette filière d'avenir.

Parce qu'en France nous avons des chercheurs, des entrepreneurs dynamiques et compétents, il y a un terreau fertile au développement de cette ressource stratégique.

Et c'est aussi à vous, femmes et hommes politiques français, d'accompagner cette évolution dans notre pays du droit et des libertés, en vous investissant de ce sujet.

Aurélia Nègre (15 janvier 2017)

Je suis très enthousiaste et pense profondément que l'intelligence artificielle pourra apporter des améliorations majeures dans la vie des gens, en particulier sur les voitures autonomes mais aussi et surtout la prédiction d'incidents/d'activité pour faciliter le travail de la police, des pompiers et des hôpitaux.

Dominique Péré, Directeur R&D aux Éditions Lefebvre-Sarrut (1^{er} décembre 2016)

L'intelligence artificielle est en passe de poser les fondements d'une nouvelle société :

- Les techniques du *machine learning* vont créer des robots logiciels qui remplaceront les tâches dites intellectuelles à faible valeur ajoutée, au même titre que les robots industriels ont fait disparaître les ouvriers spécialisés des usines.
 - Les assistants numériques
 - La classification des textes, images, vidéos, audio ...
 - Le support technique et logiciel de 1^{er} niveau
 - La publication d'informations instantanées (météos, sports, résultats d'élections...)
- De nouveaux métiers apparaissent :
- Data scientists
 - Ingénieurs spécialisés en *big data*
 - Exploitation et contrôles des données générées

De nouvelles interactions et expériences utilisateurs apparaissent :

- Systèmes conversationnels (*Chat bots*)
- Systèmes adaptatifs de préconisation (boutique Internet, claviers numériques...)

Les données deviennent un enjeu économique majeur. La plupart des gros projets d'intelligence artificielle sont portés par les GAFAMI (Google, Apple, Facebook, Amazon, Microsoft et IBM) et Baidu qui ont la particularité d'être en très grande majorité situés aux États-Unis, sauf un en Chine.

La prépondérance des données sur le logiciel est acquise, selon la logique suivante :

« En échange de l'usage de mes logiciels tu me fournis tes données que je peux utiliser à volonté ».

Pour le *machine learning* les systèmes auront de plus en plus de données d'entraînement et parce que cette technologie est accumulative, ces intelligences vont devenir de plus en plus performantes donc utiles et bien sûr utilisées.

Maintenant il se profile derrière l'intelligence artificielle classique, une technologie déjà en marche le *deep machine learning*, c'est à dire le *machine learning* entraîné dans des réseaux neuronaux.

La soudaine amélioration de Google Translate est due à la bascule vers le *deep machine learning*.

Cette technologie pose différents problèmes :

- Qui maîtrise les données d'apprentissage ?

- Comment vont évoluer ces systèmes apprenants quand ils vont apprendre avec les données produites par les systèmes de la génération précédente ?

- Quels sont les outils disponibles pour surveiller un « Neural Network » ?

- Quelles garanties aura-t-on d'un fonctionnement non biaisé et bienveillant des intelligences ?

Notre vie quotidienne va être séquencée par cette technologie, il est important que le « politique » s'empare du problème, le comprenne et agisse en conséquence. Par exemple créer une obligation aux systèmes d'apprentissage automatique d'intégrer de nouvelles données d'apprentissage, pour éviter le phénomène d'entonnoir. Avoir les moyens d'observer la qualité et la pertinence des solutions calculées.

Dans notre métier d'éditeur juridique nous voyons se profiler des « robots juges » qui vont prendre en charge les délits mineurs. Nous voyons aussi des « robots avocats » qui vont être capables de construire des raisonnements juridiques sur les cas les plus simples.

Tant que les résultats des systèmes *deep machine learning* légaux sont intermédiés par des humains, on est dans le légal assisté par ordinateur. Cependant, imaginons qu'à la suite d'un manquement grave au code de la route à 3 heures du matin, pendant que le propriétaire dort, la voiture

autonome recevant la notification du robot juge va le renvoyer vers le robot avocat du propriétaire.

Nous voyons se profiler une nécessaire éthique de l'utilisation des systèmes à apprentissage profond.

Oserais-je aborder les problèmes d'indépendances et ces systèmes fonctionnants dans la sphère militaire ?
(<https://www.youtube.com/watch?v=45-WR5RcIp0>)

Comment juger les machines ?
(http://abonnes.lemonde.fr/idees/article/2016/12/01/comment-juger-les-machines_5041562_3232.html)

Comment vont évoluer les rapports sociaux, lorsque nous serons entourés par des machines intelligentes nous assistant quotidiennement et travaillant à notre place ?

Va-t-on faire une redéfinition de la valeur travail, et de la place de l'homme au regard du travail ?

Une véritable révolution sociétale est en marche, pour ma part je souhaite que nous évitions la « *rage against the machine* » des Canuts du numérique, tout autant que l'abandon de notre destin au profit des machines, tout intelligentes soient-elles ?

Tangi Vass (10 novembre 2016)

L'enjeu à mon sens n'est pas tant les hypothétiques dangers de l'intelligence artificielle dans l'absolu et de la singularité – c'est-à-dire le moment où l'intelligence artificielle dépassera l'intelligence humaine – que la rapidité et la brutalité des impacts socio-économiques de la troisième révolution industrielle, déjà largement amorcée, dont l'intelligence artificielle est le cœur.

Le cabinet Roland Berger prévoit qu'en France « 3 millions d'emplois pourraient être détruits par la numérisation d'ici à 2025. »

« Tout l'enjeu repose donc sur la capacité de l'économie française à produire les nouvelles activités qui se substitueront à celles où les gains de productivité ont réduit le nombre d'emploi, de manière similaire à la substitution de l'industrie par les services au XX^e siècle. »

La difficulté de cette transition est accentuée par le fait que dans l'économie de la data, « *the winner takes it all* », il n'y a pas de place pour des acteurs moyens.

L'échelle nationale n'est d'ailleurs sans doute pas la bonne, comme elle ne l'était pas déjà pour l'aéronautique.

La ressource stratégique de cette 3^e révolution industrielle est la *data*, le « carburant » de l'intelligence artificielle. L'accès et le contrôle de la *data* est donc un enjeu géostratégique, comme l'étaient ceux du pétrole dans l'ère

précédente. Des cyberguerres liées au contrôle de cette ressource sont déjà en cours, des guerres traditionnelles sont possibles dans un avenir proche.

« *Nous n'avons pas de meilleurs algorithmes. Nous avons tout simplement plus de données* » Peter Norvig, directeur de recherche chez Google.

Lors du *meetup* d'hier, le représentant d'Amazon a présenté la technologie du moteur de recommandation d'Amazon, récemment ouverte en *open source*, tout en plaisantant sur le fait qu'évidemment les données d'Amazon n'étaient pas incluses. C'est un peu comme si l'Arabie Saoudite offrait gratuitement des voitures.

Les États-Unis sont aujourd'hui l'hyperpuissance de la donnée.

Citation de Stéphane Frénot (INSA Lyon) et Stéphane Grumbach (Inria) : « *Les données personnelles, tant celles directement produites par les usagers, textes, photos, vidéos, etc., que celles indirectement générées bien souvent à notre insu par les systèmes que nous utilisons, sont au cœur de l'économie de la société de l'information, au cœur, donc, de l'économie en général.* »

« *Dans ce domaine, les États européens ont fait l'impasse, ou tout au moins ont échoué à promouvoir les entreprises de la nouvelle économie. En ne construisant pas d'industrie du Web 2.0, l'Europe s'est privée de l'accès à la ressource, y compris à la ressource provenant de ses territoires. Pour des raisons historiques et politiques, l'Europe a peur des données. Elle voit dans la société de l'information une immense menace qu'il convient de circonscrire, et qui semble inhiber toute vraie ambition, sans se rendre compte que moins on développe l'industrie de l'information, plus nos données quittent le territoire.* »

« *La maîtrise de données permet également la maîtrise de certains marchés qui transitent déjà majoritairement dans certains domaines par les outils de commerce électronique américains. Faute de développer cette industrie, il est probable qu'à brève échéance nous achèterons de très nombreux biens et services produits et consommés en France, comme nos billets de train ou notre électricité, à un prestataire étranger qui dégagera une part de la valeur ajoutée et aura le contrôle de la partie la plus rentable de la chaîne industrielle, laissant les parties coûteuses comme les infrastructures à ses partenaires.* »

Charly FLORENT (29 août 2016)

L'intelligence artificielle est un domaine extrêmement fantasmé sans être réellement connu pour ses applications concrètes. Une bonne partie de mon travail au jour le jour en tant que *data scientist* dans une grande entreprise est la vulgarisation et l'évangélisation autour de ce sujet : que peut réellement faire la *data science* ?

Les machines, même si elles peuvent « apprendre », sont limitées par les algorithmes qui les animent. L'apprentissage à proprement parler est mathématique, c'est une reconnaissance de schémas, l'intelligence artificielle

ne créé pas. Ceci étant, elle peut, à partir de ce qui existe déjà, reconnaître des schémas que nous n'avions pas perçus au préalable et faire des connexions entre des éléments qui n'étaient pas évidentes. C'est pour cette raison qu'il est possible aujourd'hui de battre des champions d'échec ou de Go, pas en inventant de nouvelles stratégies mais bien en décortiquant les anciennes pour sublimer le mécanisme du jeu.

En conclusion, je pense que c'est une excellente initiative que l'intelligence artificielle soit comprise par les hautes sphères du gouvernement car elle n'est pas à négliger. Cette recherche de schémas peut s'avérer extrêmement puissante dans des domaines peu soupçonnés et apporter une amélioration nette à certains services publics (on pense déjà à la santé avec les carnets de santé virtualisés ou à la détection de fraudes fiscales).

Jean-Armand Moroni - ingénieur Centrale Paris avec une spécialisation en intelligence artificielle (28 août 2016)

L'intelligence artificielle est de nouveau sur le devant de la scène, après un moratoire de 1990 à 2005.

À l'actif :

- Des succès récents : automobile autonome, catégorisation d'image, reconnaissance de personne, IBM Watson à Jeopardy, Go...
- Une dynamique forte et relativement ouverte : communauté de chercheurs, d'amateurs, de sociétés (les grands du Web), de sites (Kaggle pour le *data mining*), mouvement *open data*, etc. L'intelligence collective du Web.
- Des données en quantité suffisante, des algorithmes pour les traiter (en particulier le *deep learning*), des puissances de calcul qui continuent de croître, des outils *open source* dédiés (bases NoSQL, bibliothèques ex. TensorFlow).
- Des robots avec des capacités physiques qui s'améliorent (même si toujours loin de celles des humains) ; des plateformes accessibles (Arduino, Raspberry, capteurs et actionneurs) et l'impression 3D simplifient l'expérimentation.

Au passif :

- Toujours pas de conversation possible avec une intelligence artificielle. Toujours pas de sens commun. Les succès remportés ne vont pas au-delà de ce que sait faire un animal, qui n'a pourtant pas accès au langage structuré. Risque d'atteindre le plus haut du « *hype cycle* » et d'en dégringoler ensuite : un troisième « hiver de l'IA ».

- Les algorithmes d'apprentissage actuels semblent beaucoup moins efficaces que ceux mis en œuvre par un humain (ex. la quantité de données nécessaires) : est-on sur la bonne piste ?
- Les *open data* sont hétérogènes ; l'absence de standardisation complique leur croisement ; le *semantic Web* patine.
- Les retombées commerciales du *big data* pour les sociétés qui y investissent ne sont pas évidentes, en dehors de cas particuliers type Google.
- Très forte domination des États-Unis pour la mise en œuvre, même si la recherche est beaucoup mieux distribuée (ex. Yann LeCun, père français du *deep learning*).

ANNEXE 7 : COMPTE RENDU DE L'AUDITION PUBLIQUE DU 19 JANVIER 2017¹

I. INTRODUCTION

1. M. Bruno Sido, sénateur, premier vice-président de l'OPECST

Monsieur le président, mesdames, messieurs les députés, mesdames, messieurs les sénateurs, mesdames, messieurs les professeurs, en ma qualité de sénateur et de premier vice-président de l'Office parlementaire d'évaluation des choix scientifiques et technologiques, j'ai le plaisir de vous accueillir ce matin au Palais du Luxembourg pour une journée d'auditions publiques sur l'intelligence artificielle (IA).

Le président de l'OPECST, M. Jean-Yves Le Déaut, député, vous présentera en détail l'originalité de cette délégation parlementaire bicamérale. Je me contenterai de saluer le travail entrepris depuis quelques mois, et à un rythme soutenu, par la sénatrice Dominique Gillot et le député Claude de Ganay, les deux rapporteurs désignés par l'Office pour répondre à la saisine émanant de la commission des affaires économiques du Sénat sur l'intelligence artificielle.

D'ailleurs, il est à noter que Mme Dominique Gillot espérait depuis plusieurs années travailler sur ce thème. C'est dire que l'Office ne découvre pas aujourd'hui le sujet de l'intelligence artificielle, d'autant que cette audition publique a été précédée de plusieurs autres études ou auditions publiques abordant déjà ce thème, sous un angle ou un autre. Je pense, par exemple, à la récente audition publique sur les robots et la loi, mais aussi à celle sur le numérique au service de la santé ou encore à celle sur l'agriculture et le traitement des données massives.

En outre, j'ai eu l'honneur d'être co-rapporteur de l'étude relative à la sécurité numérique, qui pourrait être beaucoup mieux garantie si les recommandations préconisées par l'Office étaient suivies, notamment celles qui constituent son *vade-mecum* pour la sécurité numérique des entreprises.

¹ Ces auditions ont constitué un record d'audience pour le Sénat selon les données disponibles à la fin du mois de janvier 2017 : sur Facebook, la vidéo a été regardée en direct par 97 182 personnes, puis en différé par 8 600 personnes. Sur le site Internet du Sénat, le nombre de consultations de la vidéo s'élève à 1 839.

La chaîne Public Sénat en a conçu une émission spéciale, consultable sur : <https://www.publicsenat.fr/emission/les-matins-du-senat/table-ronde-intelligence-artificielle-51796>

L'intelligence artificielle a donc déjà piqué la curiosité des membres de l'Office, qui ne manqueront pas de regarder avec intérêt la vidéo de la présente audition sur le site du Sénat.

Quant à la publicité ultérieure des travaux de ce jour et de leurs retombées positives dans l'étude à paraître, il faut savoir que chaque étude de l'Office, incluant les contributions des intervenants sous forme de comptes rendus, donne lieu à la présentation de ses conclusions devant la commission qui l'a saisi puis, devant le Sénat, à l'occasion d'un débat en séance plénière auquel participe le ministre compétent : Mme Axelle Lemaire, secrétaire d'État chargée du numérique et de l'innovation, nous fera d'ailleurs l'honneur d'être parmi nous cet après-midi.

Avant de céder la parole au président de l'Office, je souhaiterais également souligner que l'OPECST ne se contente pas de travailler le plus en amont possible du sujet qu'il traite, mais qu'il essaie également de bien orienter sa réflexion vers les aspects les moins bien explorés.

En l'occurrence, il ne s'agit pas, pour nous, de dresser un panorama de l'historique des différentes vagues d'intérêt suscitées depuis un demi-siècle par l'intelligence artificielle. Elles sont présentées, à chaque occasion, comme une révolution décisive devant tout emporter sur son passage, ce qui a été largement démenti à plusieurs reprises.

Cette fois-ci, il semblerait que divers éléments se conjuguent pour que la prochaine vague soit plus puissante que les précédentes et induise un changement de société qui pourrait dépasser nos perspectives actuelles d'imagination.

Dans ce contexte, il me semble que les rapporteurs de l'Office ont fort opportunément choisi de dresser une cartographie mettant plus particulièrement en valeur un aspect original, à savoir l'approche éthique des problèmes posés par l'émergence accélérée de l'intelligence artificielle, sans oublier les interconnexions ultérieures qui ne manqueront pas de se nouer avec des domaines inattendus. Mais je n'en dirai pas plus pour laisser à la fois au président et aux deux rapporteurs le plaisir et l'honneur de vous présenter l'Office et de vous exposer leur démarche.

Je vous remercie de nouveau de vous être déplacés et d'avoir mobilisé toutes vos capacités cognitives pour rendre notre intelligence du sujet la moins artificielle possible.

2. M. Jean-Yves Le Déaut, député, président de l'OPECST

Mesdames, messieurs, je vous remercie également d'être venus ce matin. Le sujet que nous allons aborder est en prise directe avec une actualité brûlante. L'impact du développement de l'intelligence artificielle sur nos sociétés est au cœur des discussions du forum de Davos ; ce thème récurrent

est évoqué depuis plusieurs années. Mme Mady Delvaux, que nous recevrons cet après-midi, a publié, au titre d'un groupe de travail de la commission des affaires juridiques du Parlement européen, un rapport sur les règles de droit civil concernant la robotique et l'intelligence artificielle, dont les conclusions donnent lieu à discussion dans la mesure où est, notamment, posée la question de la responsabilité des robots.

Moi-même, en tant que rapporteur général des Sciences et des technologies de l'Assemblée parlementaire du Conseil de l'Europe, je prépare un rapport sur les convergences technologiques, l'intelligence artificielle et les droits de l'homme, en le ciblant plus particulièrement sur les problèmes que pose l'intelligence artificielle au regard des droits de l'homme. Plusieurs rapports vont être publiés sur le sujet. Aussi, il est opportun que l'Office s'en soit saisi, en chargeant Dominique Gillot et Claude de Ganay de réaliser cette étude.

Notre manière de travailler est originale - elle a de plus en plus cours à l'Assemblée nationale et au Sénat - en ce qu'elle consiste à organiser des auditions publiques collectives contradictoires où interviennent des parlementaires et des représentants de la société civile, qu'il s'agisse d'universitaires, de chercheurs ou de responsables d'associations. Ces discussions, au cours desquelles des avis différents s'expriment, nous permettent d'envisager des évolutions législatives. L'Office travaille en amont de la législation. Sur des sujets complexes, tels que celui de l'intelligence artificielle, il est évident que les parlementaires ne sont pas armés de la même manière pour les décortiquer. Or légiférer exige de bien comprendre les sujets.

De manière très simpliste, certains déplorent l'absence de parlementaires dans les hémicycles, mais, pour débattre d'un sujet et légiférer, il faut se l'être approprié. Or c'est compliqué. Tel est le rôle de l'Office. D'ailleurs, de plus en plus de collègues nous posent des questions simples. On nous a récemment interrogés sur le compteur *Linky* par exemple ; on nous a aussi demandé pourquoi on ne développe pas des vignes résistantes au mildiou ou à l'oïdium. On examine ces questions et on essaie de trouver des solutions.

L'intelligence artificielle est un sujet majeur, et il sera très difficile de régler la question de la législation à l'échelon national - une proposition va nous être transmise à l'échelon européen. Il convient à l'évidence de traiter nombre de ces sujets à l'échelle internationale. Dès lors que le robot devient intelligent, que la frontière entre l'homme et la machine va s'estomper ou, plutôt, que l'homme et la machine vont continuer à travailler ensemble, qui sera responsable ? Le problème de la responsabilité et de la réparation des dommages se pose. On parle souvent de robots automatisés dans le domaine de la guerre, mais ce ne sont pas les seuls - heureusement ! L'audition publique sur les robots et la loi, que nous avons organisée en décembre 2015 et qui était préparatoire à celle d'aujourd'hui, a prouvé la pertinence d'aménager les règles du droit pour tenir compte de nouvelles formes de

responsabilité, avec le déploiement des robots induits, plutôt que de créer un statut et des régimes de responsabilité spécifiques avec une personnalité juridique pour les robots, une solution à laquelle je n'adhère pas personnellement. Mes collègues rapporteurs en parleront.

Dans le domaine de la médecine notamment, la frontière s'estompe entre l'homme et la machine : on est passé de l'homme soigné à l'homme réparé, mais va-t-on passer à l'homme augmenté ? Certains en parlent. Cette seule question pose des problèmes éthiques. La loi doit permettre de résister à des pressions ou à des contraintes, qui imposeraient à des individus, par exemple, de se soumettre à des technologies de nature à améliorer leurs performances ; je pense aux domaines du sport, du jeu, voire à celui du travail. Toutes ces questions doivent donc être posées.

Ces nouvelles technologies induisent déjà des changements de grande ampleur. La frontière entre le médical et le non médical s'atténue également, et il en est de même pour le naturel et l'artificiel. Avec la biologie de synthèse, on recrée le vivant à partir de molécules recomposées. Pour travailler actuellement à l'élaboration d'un rapport sur le *genome editing*, je puis vous dire que l'on voit aujourd'hui les grands progrès réalisés en matière de modification ciblée du génome, avec un abaissement des coûts. Cela pose aussi un certain nombre de problèmes.

La question de l'équilibre entre les innovations, qui sont nécessaires, et les droits de l'homme est donc posée. D'autres problèmes se posent, tels que les fractures numériques ou l'accès inégal à Internet ; des fractures mondiales assez criantes sont également à noter. Le professeur Jan Helge Solbakk du Centre d'éthique médicale à la faculté de médecine de l'Université d'Oslo résume bien les choses : le progrès n'est pas synonyme d'accessibilité universelle. L'Office en est d'accord, même s'il faut continuer à employer le mot « progrès », pour lequel nous nous battons, il doit être maîtrisé et partagé. C'est l'un des points sur lesquels vous allez vous exprimer.

En conclusion, j'observe que, au travers de l'intelligence artificielle et de la convergence des technologies NBIC, les nanotechnologies, biotechnologies, informatique et sciences cognitives - la fameuse convergence entre le neuro, le bio, l'info et le cogno -, l'homme va pouvoir augmenter ses capacités grâce aux machines. Toutefois, comme je l'ai souligné précédemment, l'homme augmenté risque de succéder à l'homme soigné : on sait aujourd'hui fabriquer des interfaces fonctionnelles entre le cerveau et l'ordinateur.

Toutes ces questions méritent une réflexion prospective, que nos deux collègues ont souhaité mener. Restons modestes dans un premier temps : ne pensons pas résoudre cette question en pleine évolution et en pleine mutation. Mais les échanges avec des spécialistes et des parlementaires participent d'une bonne méthode de travail. Aussi, je suis très heureux que l'OPECST puisse apporter sa contribution, en vue d'adapter

peut-être le cadre législatif français. Les recommandations de nos deux rapporteurs seront débattues au mois de mars prochain ; il s'agira d'ailleurs du dernier rapport de l'Office pendant cette législature. Au titre de nos travaux, nous débattons, la semaine précédente, du rapport relatif au *genome editing* et de l'évaluation de la stratégie nationale de recherche en février. Par ailleurs, une audition publique sur le compteur *Linky* nous a été demandée par nos collègues.

II. PREMIÈRE TABLE RONDE, PRÉSIDIÉE PAR M. CLAUDE DE GANAY, RAPPORTEUR : LES TECHNOLOGIES RELEVANT DE L'INTELLIGENCE ARTIFICIELLE

1. Mme Dominique Gillot, sénatrice, membre de l'OPECST, rapporteure

Comme vient de le souligner le président Le Déaut, nous avons l'ambition forte d'aider la décision politique en l'appuyant sur la réflexion scientifique et le dialogue entre les scientifiques et les politiques. Cette gouvernance, souvent réclamée et exaltée, est quelquefois difficile à mettre en place du fait de la frontière qui existe entre ces deux mondes. L'OPECST a vraiment vocation à éclairer les parlementaires et, au-delà, les décideurs de la puissance publique. Tel est notre objectif avec le rapport de l'Office sur l'intelligence artificielle, afin de participer à la formation de la décision politique.

Pour ma part, je modérerai quelque peu les propos de mon collègue : l'intelligence artificielle ne signifie pas seulement que l'homme doive se soumettre, car il peut aussi en bénéficier, profiter de nouvelles opportunités. De même, il ne s'agit pas simplement d'être assujéti, on peut aussi maîtriser pour partager des progrès. J'ai une vision plutôt optimiste de l'avancée de la science pour autant que l'homme se donne les moyens de continuer de l'accompagner et de la maîtriser. Le meilleur rendement, c'est l'homme et la machine ; ce n'est pas la machine seule.

2. M. Claude de Ganay, député, membre de l'OPECST, rapporteur

Dominique Gillot vient de préciser le cadre de travail de l'Office.

La première table ronde que j'ai l'honneur de présider porte sur les technologies relevant de l'intelligence artificielle. Nous avons choisi ce titre pour évoquer le caractère multiple de ces technologies et préciser tout particulièrement ce que l'on met derrière le label « intelligence artificielle ».

Je demanderai aux intervenants de respecter un temps de parole de sept ou huit minutes maximum, afin d'avoir un temps raisonnable pour notre débat interactif.

Je laisse tout d'abord la parole à Jean-Gabriel Ganascia, professeur à l'Université Pierre-et-Marie-Curie Paris-VI, rattaché au laboratoire d'informatique de l'Université Paris-VI. Vous présidez depuis quelques mois le comité d'éthique du CNRS. Physicien et philosophe d'abord, puis informaticien, passionné par les sciences cognitives, vous êtes un spécialiste reconnu de l'intelligence artificielle. Vous avez d'ailleurs publié de nombreux ouvrages à ce sujet : le prochain, qui devrait être publié en février, devrait s'intituler : « *Le mythe de la singularité. Faut-il craindre l'intelligence artificielle ?* » Nous allons vous écouter afin de répondre de manière argumentée à cette question.

3. M. Jean-Gabriel Ganascia, professeur à l'Université Pierre-et-Marie-Curie, Paris-VI

Monsieur le président, mesdames, messieurs les sénateurs, mesdames, messieurs les députés, je vous remercie tout d'abord de m'avoir invité. Pour respecter scrupuleusement le temps qui m'a été alloué, j'entre dans le vif du sujet.

J'ai intitulé mon exposé sur l'intelligence artificielle « *Naissance et renaissances de l'intelligence artificielle* ». Comme l'a relevé précédemment M. Sido, il existe différentes renaissances de l'intelligence artificielle. D'une certaine façon, il y a un caractère cyclique, que j'appelle « un temps tressé ». Les technologies de l'information en général font que des technologies oubliées redeviennent importantes.

L'autre élément important tient à l'article défini singulier : l'intelligence artificielle. Il s'agit d'une discipline scientifique ; on ne peut donc pas parler d'une intelligence artificielle. Cela substantifierait les objets mus par l'intelligence artificielle, ce qui, à mon sens, est incorrect.

Permettez-moi de faire une petite rétrospective. L'idée de faire une machine qui raisonne est très ancienne : cela remonte à Leibniz. Au XIX^e siècle, Jevons a utilisé l'algèbre de Boole pour faire un piano mécanique, qui faisait du raisonnement de façon automatique. Ensuite, en 1943, avant la construction du premier ordinateur électronique, la cybernétique - ce nom ne sera donné qu'en 1946 - s'est attachée à la simulation du raisonnement sur des machines.

Sans entrer dans le détail, deux articles princeps paraissent en 1943, un article sur la notion de rétroaction et un second sur la notion de réseaux de neurones formels, une idée introduite par deux personnes, Warren McCulloch et Walter Pitts, un jeune mathématicien alors âgé de vingt ans. Ce

dernier montre que des réseaux très simples, avec de petits automates organisés sur trois couches connectées entre elles, permettent de réaliser n'importe quelle fonction logique. Il établit alors un parallèle entre ce que l'on connaissait à l'époque du cerveau et l'ingénierie, cette fabrication d'outils. Toutefois, il faut réussir à connecter les trois couches du réseau entre elles, et l'apprentissage de ces connexions fera l'objet des travaux menés dans les soixante ou soixante-dix années qui suivent.

La vraie naissance de l'intelligence artificielle date de 1956 : un programme de recherche a été écrit par deux jeunes mathématiciens âgés de vingt-huit ans, John McCarthy et Marvin Minsky. Ce programme de recherche est primordial - c'est l'équivalent de ce que fut, en son temps, la déclaration de Galilée au début de l'âge moderne lorsqu'il affirme que la nature s'écrit en langage mathématique : l'esprit peut se décomposer et se simuler avec une machine. Je l'ai souligné dans ma diapositive, tous les aspects de l'apprentissage et toutes les autres caractéristiques de l'intelligence peuvent être décrits de façon si simple qu'une machine peut les simuler. C'est ce coup d'envoi épistémologique de l'intelligence artificielle qui est central pour l'âge moderne.

Quelles sont les différences caractéristiques de l'intelligence ? J'ai utilisé une décomposition classique qu'utilisent les spécialistes des sciences cognitives, à savoir cinq fonctions cognitives classiques : les fonctions réceptives, avec l'analyse des formes, des images, des sons, etc. ; la mémoire et l'apprentissage ; les questions de raisonnement et de pensée, qui sont simulées par les machines ; les fonctions expressives - comment parler ? Comment échanger ? - et les fonctions exécutives - les robots dont on a parlé tout à l'heure.

Je parlerai maintenant des renaissances. Quelles sont les techniques ? Elles sont nombreuses. Mais je me contenterai de vous parler de l'apprentissage profond, le *deep learning*. Le cycle des renaissances commence en 1943 avec les réseaux de neurones formels : un réseau à trois couches peut réaliser n'importe quelle fonction logique : c'est une propriété d'universalité des réseaux. Le problème réside dans le fait d'établir des connexions entre la première et la deuxième couche, et la deuxième et la troisième. Ce travail de programmation étant extrêmement fastidieux, on essaie de le faire automatiquement. Marvin Minsky avait essayé de faire ce travail dans sa thèse, sans y parvenir. En 1959, Frank Rosenblatt avait inventé le perceptron, mais ce système n'apprenait les connexions que sur le réseau à deux couches, ce qui était très pauvre.

La nouvelle renaissance intervient en 1986, avec des théories mathématiques nouvelles, la *Back Prop*, la rétropropagation de gradient permet d'assurer l'apprentissage. Des mathématiciens essaient de formaliser tout ce travail avec d'autres techniques. À partir de 2010, on réutilise les réseaux de neurones non plus à trois, mais à treize couches, mais je n'entrerai pas dans le détail. Les résultats sont époustouffants, notamment pour la reconnaissance des formes. Je citerai un exemple marquant, la

reconnaissance des visages. Différentes sociétés s'intéressent à ce sujet, la société *FaceNet* de Google ou *DeepFace* de Facebook entre autres : pour la première, le taux de reconnaissance est de 99,63 %, mais il faut apprendre sur 200 millions d'images, ce qui pose un certain nombre de questions, que nous aborderons ultérieurement.

Autre enjeu majeur, voilà moins d'un an, la machine AlphaGo a battu celui que l'on considérait à l'époque comme le meilleur joueur de Go au monde. Certes, il s'agit d'un jeu, mais celui-ci est extrêmement sérieux. Du point de vue scientifique, John McCarthy disait des jeux qu'ils étaient la drosophile de l'intelligence artificielle. C'est en fait ce qui va permettre de tester l'ensemble des techniques, et cela fait appel à de l'apprentissage profond, ainsi qu'à d'autres techniques d'apprentissage, tel l'apprentissage par renforcement.

Pour clore cette présentation, je parlerai des enjeux liés à l'intelligence artificielle : les enjeux éthiques certes, dont nous parlerons ultérieurement, mais aussi les enjeux économiques, qui sont majeurs.

Je citerai une technique qui a suivi toute l'histoire de l'intelligence artificielle, celle des agents conversationnels. Le test de Turing en 1950 imaginait une conversation entre une machine et un homme, la confusion qu'il peut y avoir entre un homme et une machine. Dans les années soixante, Joseph Weizenbaum a aussi imaginé un agent conversationnel. Depuis quelques années, de nombreux travaux ont été publiés sur ce sujet. L'agent conversationnel Tay a défrayé la chronique l'an dernier ; la société Amazon a développé le système Écho, Google l'Assistant, et Apple a fait part de son intention de développer un assistant conversationnel nommé Viv, qui doit prendre la relève de Siri. Ils pensent qu'ils vont constituer le segment terminal dans toute une chaîne de valeurs. La publicité sera « *cool* » avec les agents conversationnels : vous pourrez commander une pizza directement de votre bureau. L'enjeu est simple : les personnes qui maîtriseront la technologie de cet agent auront alors un rôle déterminant sur toute la restauration rapide.

J'évoquerai enfin les enjeux politiques, les enjeux de souveraineté.

Si l'on considère les fonctions régaliennes de l'État, à savoir la défense, la sécurité intérieure, les finances et la justice, on voit qu'elles sont toutes transformées par l'intelligence artificielle, car de nouveaux acteurs interviennent.

Pour citer la justice, j'ai été surpris de constater que les États-Unis utilisent des systèmes prédictifs ; cela n'existe pas encore en France. Le site *crowdlaw.org* peut faire de la législation collective en s'adressant à la population. Tout cela fait réfléchir.

Concernant la finance, les États avaient le privilège de battre monnaie ; c'est un peu moins vrai avec l'Europe et cela l'est encore moins avec les monnaies virtuelles.

S'agissant de la sécurité intérieure, dans le cadre de leur programme d'exemption de visas, les États-Unis exigent depuis quelques semaines que soient notés sur le formulaire *Esta* les pseudonymes sur les réseaux sociaux. Cela signifie, d'une certaine façon, que l'état civil est mieux défini par ces réseaux que par l'État. On peut aussi parler de la reconnaissance des visages, qui assure une certaine sécurité. Là encore, l'État ne peut y procéder parce qu'il ne dispose pas des images et n'a pas toutes les capacités pour le faire, même d'un point de vue légal.

Enfin, pour ce qui concerne la défense, dont je croyais qu'elle restait le privilège de l'État, j'ai appris voilà quelques semaines que les grandes entreprises veulent être autorisées à avoir une attitude non seulement défensive sur le cyberspace, mais également offensive, pour se battre contre leurs agresseurs.

M. Claude de Ganay, rapporteur. - Monsieur Gérard Sabah, vous êtes directeur de recherche honoraire au CNRS. Vous avez longtemps travaillé au sein du laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur à Orsay, le LIMSI. En tant que membre titulaire de l'académie des technologies, vous avez produit en 2009 une brochure sur l'intelligence artificielle, qui faisait un point très intéressant sur l'histoire de ces technologies. Je vous demande donc de bien vouloir nous dresser le bilan de ces technologies.

4. M. Gérard Sabah, directeur de recherche honoraire au CNRS

Monsieur le président, mesdames, messieurs les sénateurs, mesdames, messieurs les députés, je vous remercie de cette invitation. Je suis content de constater que l'audition est ouverte à la presse, même si elle est assez peu représentée : elle sera ainsi informée de nos travaux. Cela lui permettra d'éviter de publier des articles intitulés « *Il n'y aura jamais d'intelligence artificielle. L'idée d'intelligence artificielle présuppose qu'il n'existe qu'une seule forme d'intelligence.* » Bien sûr, cela n'est pas le cas. Nous avons essayé, dans le cadre des travaux relatifs à l'intelligence artificielle, d'identifier un certain nombre de caractéristiques de cette intelligence. Pour les premières d'entre elles que j'ai listées sur ma diapositive, l'informatique, en général, et l'intelligence artificielle, en particulier, sont capables de réaliser la plupart des éléments notés, et les recherches pointues portent sur la réalisation des derniers éléments de la liste.

Je rappelle très brièvement les principaux outils de l'intelligence artificielle classiques : au tout début, le développement d'arborescences, la recherche dans les arbres, avec des procédures d'amélioration pour accélérer les choses ; le système expert dans les années quatre-vingt ; le développement de logiques non classiques pour essayer de rendre compte de

la complexité des raisonnements humains ; la programmation par contrainte, les réseaux de neurones, les algorithmes génétiques, raisonnement par analogie. On le voit bien, différents modes de raisonnement ont été mis en œuvre. Les mécanismes de l'intelligence artificielle sont également très variés.

Trois courants de pensée relativement différents sous-tendent l'intelligence artificielle : l'analogie symbolique, dont Jean-Gabriel Ganascia a parlé, est fondatrice de la discipline – les processus mentaux sont censés se résumer à des manipulations de symboles ; l'intelligence artificielle distribuée, où la pensée est vue comme un phénomène collectif – le développement de nombreux agents élémentaires peut être synthétisé et produire un comportement intelligent – et la métaphore des réseaux où l'intelligence est considérée comme une diffusion d'activités au niveau des réseaux de neurones.

Je vous ferai part de quelques succès : les échecs, le backgammon, le jeu de Go où le champion du monde a été battu, le développement d'interfaces vocales, et Watson, qui a battu les meilleurs joueurs mondiaux au jeu Jeopardy.

À cet égard, je souligne que Watson disposait de quinze téraoctets de mémoire vive et utilisait près de trois mille processeurs. Cet ordinateur est donc loin d'être similaire à ceux dont nous disposons. IBM a dit que, si le programme avait été mis sur un ordinateur de bureau, les logiciels auraient mis deux à trois heures pour répondre à une question. Le jeu consistait à trouver une question dont la réponse lui avait été communiquée. À la question « *Un écrivain dont le héros est Hercule Poirot* », il fallait répondre : « *Qui est Agatha Christie ?* » Watson disposait de deux cents millions de pages d'encyclopédies, de dictionnaires, de livres, d'articles de journaux, de scénarios de films, etc. Comme ses adversaires, il n'avait pas accès à Internet, mais il avait toutes ces données en mémoire.

Actuellement, l'apprentissage profond lié aux données massives est à la mode. On dispose de quantités de données absolument extraordinaires en la matière. Comme l'a montré Jean-Gabriel Ganascia, l'apprentissage profond, qui consiste à utiliser de nombreuses couches intermédiaires, permet de mettre en évidence un certain nombre de concepts élémentaires utiles pour la reconnaissance d'images ou la compréhension de textes, par exemple. Cela a été rendu possible par le développement de la puissance de calculs des machines et par l'utilisation de nouvelles puces graphiques, utilisées aussi au niveau du calcul.

Quelles sont les limites ?

Dans ce type de mécanisme, le réseau produit effectivement quelque chose, mais n'a pas de possibilité d'explication. On ne comprend pas forcément pourquoi cela marche. Cela peut être opposé aux travaux relatifs aux mécanismes d'apprentissage à partir de très peu d'exemples. En 2015,

dans la revue *Science*, un article a montré comment on pouvait réaliser des apprentissages à partir d'un ou deux exemples.

Le programme *DeepDream* créé par Google reconnaît, disaient-ils, une image de chat après plusieurs centaines de milliers ou de millions d'analyses d'images. Que veut dire reconnaître un chat ? Il ne faut pas imaginer que l'ordinateur a compris ce qu'était un chat ou qu'il a mis en évidence ce concept. En fait, il met dans la même classe toutes les images de chat. Si l'on demande à l'ordinateur d'autres caractéristiques du chat, il ne pourra évidemment pas y répondre.

Je formulerai quelques réflexions futures, en insistant sur deux points.

La question de l'hybridation. Le fantasme actuel de l'auto-apprentissage ? Actuellement, on a l'impression que l'hybridation va tout résoudre, mais l'intelligence artificielle ne se résume pas à l'apprentissage machine, en particulier aux techniques de réseaux de neurones profonds, d'apprentissage profond.

L'intelligence artificielle, comme prothèse, vise à ajouter aux compétences humaines et conduit à se demander dans quelle mesure l'homme va pouvoir éventuellement être « augmenté » par les techniques de la machine.

Plus loin, il y a la notion de conscience, avec la capacité à se représenter soi-même et à raisonner sur ses représentations : on pourrait se demander dans quelle mesure les machines pourraient acquérir ce type de capacités.

J'ai essayé de résumer les avantages et les inconvénients des deux approches des systèmes symboliques de l'intelligence artificielle classique et les réseaux de neurones profonds : difficultés à généraliser pour les systèmes symboliques, généralisation naturelle pour les réseaux de neurones, etc. Essayer de faire collaborer ces deux approches présente manifestement un certain nombre d'avantages, dont certains peuvent être regroupés.

Je ne m'attarderai pas sur l'hybridation entre l'homme et l'intelligence artificielle, car Laurent Alexandre, qui doit intervenir cet après-midi, en parlera beaucoup mieux que moi.

Les hybridations entre les systèmes symboliques et les systèmes neuronaux : le but est d'obtenir de meilleures performances globales, comme je l'ai dit précédemment, avec des champs d'application plus larges, afin de permettre l'acquisition de nouvelles connaissances à partir de sources différentes et pas toujours cohérentes.

Comment y parvenir ?

Il faut, par exemple, essayer d'adapter la structure du réseau avec des algorithmes génétiques ; introduire des modules symboliques dans les réseaux, afin de permettre l'utilisation de logiques floues, de systèmes de

règles. Diverses questions se posent, que je n'ai pas le temps d'évoquer ici : comment intégrer ces différents modules ? Comment interagissent-ils entre eux ?

Par ailleurs, concernant la conscience : selon Gerald Edelman, prix Nobel en neurobiologie, les fonctionnalités nécessaires à une véritable intelligence sont celles qui sont fondées sur l'inconscient et qui permettent l'émergence de la conscience chez l'homme. Pourquoi cette assertion ne serait-elle pas vraie pour les machines ?

Souvent, la première question que l'on peut se poser est la suivante : un robot ou un système pourra-t-il être conscient ou peut-on reproduire la conscience humaine dans une machine ? À mon sens, ce n'est pas la bonne question. Il faut se demander : parmi les fonctionnalités que l'homme attribue à sa conscience, lesquelles pourront-elles être mises en œuvre dans des futurs robots ? La notion de conscience est un concept si multiforme qu'il n'est pas possible d'en donner une définition précise. L'idée est donc d'imaginer les différentes fonctionnalités liées à cette notion. L'intelligence artificielle ne cherche pas à reproduire un être humain ; il existe d'autres moyens...

Dans le cadre d'un groupe de travail que j'ai animé au sein de l'académie des technologies, nous avons essayé d'identifier dans notre rapport ces différentes fonctionnalités. Sur ma diapositive, j'ai indiqué en vert les différentes fonctionnalités que l'on sait déjà mettre en œuvre ; en orange, celles pour lesquelles on commence à avoir quelques idées, que l'on sait presque faire ou que l'on pourra faire bientôt et, en rouge, celles qui semblent en dehors de nos compétences actuelles. On le voit, un certain nombre de fonctionnalités sont réalisables dès maintenant ou le seront bientôt.

Que faut-il de plus pour avoir une intelligence artificielle forte ?

Il faudrait que ces mécanismes soient capables de prendre du recul, de changer de mode de raisonnement, de trouver des analogies entre des univers différents, de développer des théories, de les vérifier par l'expérimentation et d'avoir envie de faire quelque chose. Qu'est-ce que cette notion d'envie, qui peut être une motivation pour agir ?

Dans un monde ouvert, se posent des problèmes de relations. Dans un monde fermé, on sait préciser les limites de ces techniques. Il faut également tenir compte de la vision stratégique à long terme des « GAFA » – Google, Apple, Facebook et Amazon –, qui semblent pour l'instant prendre le contrôle sur tous ces aspects. Il importe que l'Europe se positionne en la matière.

Alors sur cette intelligence artificielle forte : est-ce de la science-fiction ou n'est-ce qu'une question de calendrier ? Il n'y a jamais eu d'impossibilité prouvée scientifiquement, ni de preuve de possibilité non plus. Toutefois, il faut le souligner, l'intelligence artificielle n'est pas une

technique comme les autres au sens où elle débouche sur une certaine autonomie et qu'elle a des possibilités d'apprentissage, d'auto-modifications, etc. Le problème fondamental qui se pose concerne la validation : comment peut-on garantir les propriétés d'un tel système ? Et comment parvenir à conserver le contrôle face aux robots éventuellement dotés de ces capacités ?

M. Claude de Ganay, rapporteur. – Je vous remercie pour ce bilan.

Monsieur Yves Demazeau, vous êtes directeur de recherche au CNRS à Grenoble. Mais vous êtes surtout président de l'Association française pour l'intelligence artificielle, et c'est à ce titre que nous vous avons invité. Vous allez pouvoir nous dresser un tableau de la recherche française dans le domaine de l'intelligence artificielle.

5. M. Yves Demazeau, président de l'Association française pour l'intelligence artificielle

Monsieur le président, mesdames, messieurs les sénateurs, mesdames, messieurs les députés, je vous remercie de votre invitation. À vrai dire, je suis d'abord directeur de recherche au CNRS et, pendant mon temps libre, président de l'Association française pour l'intelligence artificielle (AFIA).

Je vous présenterai l'AFIA et vous exposerai les trois objectifs que nous recherchons depuis 2011, en essayant de dresser un panorama de l'activité de recherche dans le domaine de l'intelligence artificielle en France, du point de vue des chercheurs.

L'intelligence artificielle, discipline née il y a 60 ans, existe donc depuis longtemps. Il s'agit d'une force tranquille qui continue de progresser.

Elle s'intéresse à l'imitation du comportement humain, par l'usage de l'informatique. L'objectif est de créer des systèmes dont le comportement s'apparente à celui de l'être humain.

Les thèmes scientifiques développés à la dernière *International Joint Conference on Artificial Intelligence* (IJCAI) qui s'est tenue à New York en juillet 2016, étaient les systèmes multi-agents, l'intelligence artificielle et le web, l'analyse combinatoire et la recherche heuristique, l'apprentissage automatique, la planification, le contexte humain, et d'autres encore. Lorsque j'ai commencé à travailler sur ces sujets, il y a près de 35 ans, l'intelligence artificielle regroupait tous ces thèmes au sein de cette même conférence IJCAI. L'enjeu était alors surtout de comprendre comment construire une machine intelligente composée de tous ces domaines. Cette question s'étant révélée très complexe, le problème a été décomposé. Et si la lumière est projetée de temps en temps sur tel ou tel domaine, l'enjeu ultime reste l'intégration de ces différentes parties.

La recherche en intelligence artificielle, comme toutes les autres recherches, suit des cycles, du fait des découvertes et des couvertures médiatiques. Ainsi, on a beaucoup parlé à un moment donné d'intelligence artificielle à propos des systèmes experts, on parle beaucoup actuellement d'apprentissage profond. Ce focus actuel est dû à des résultats spectaculaires, qui ne sont possibles maintenant que par la conjonction de la disponibilité de masses de données, et de la capacité de calcul des machines qui ne fait que croître. Si ces progrès n'avaient pas existé, le *buzz* actuel pourrait tout aussi bien concerner plutôt l'autonomie, en particulier en ce qui concerne les véhicules autonomes. Alors, sur quoi le focus sera-t-il mis dans quelques années ?

En France, l'AFIA a identifié dix thèmes de recherche, organisés en à peu près autant de communautés, sinon de conférences :

- apprentissage automatique ;
- extraction et gestion des connaissances ;
- interaction avec l'humain ;
- reconnaissance des formes, vision ;
- représentation et raisonnement ;
- robotique et automatique ;
- satisfaisabilité et contraintes ;
- sciences de l'ingénierie des connaissances ;
- systèmes multi-agents et agents autonomes ;
- traitement automatique des langues.

Au passage, l'apprentissage profond, qui a réveillé l'intelligence artificielle ces dernières années, ne correspond selon nous qu'à 50 % de l'apprentissage automatique, qui est lui-même l'un de ces dix thèmes. Il y a donc aussi tout le reste... Cet éclatement en sous-disciplines, qui se sont forgées chacune une identité ces trente dernières années nous pénalise aujourd'hui, et nous avons donc tout intérêt à les faire interagir entre elles.

Certes, la vision peut paraître un peu académique, mais 95 % des membres de l'AFIA sont issus du monde académique, contre trop peu d'entrepreneurs et d'industriels encore. D'où vient ce décalage ?

Depuis soixante ans, les États-Unis ont bien compris leur intérêt stratégique en matière d'intelligence artificielle. Créateurs de la discipline - au moins du point de vue occidental - ils nous imposent leurs vues et pillent nos forces françaises en intelligence artificielle. C'est très visible dans les conférences internationales comme IJCAI, mais c'est aussi semble-t-il le cas ici. Les industriels invités à la présente audition l'illustrent également : je vois surtout des représentants des GAFAMI.

Suivant les acceptions, l'intelligence artificielle peut représenter jusque 50 % de l'informatique. Ainsi, la moitié des *Lecture Notes in Computer Science* (LNCS) aux éditions *Springer-Verlag* est consacrée à l'intelligence artificielle - (*Lecture Notes in Artificial Intelligence* ou LNAI)-. En France, l'intelligence artificielle n'atteint certes pas ce niveau de reconnaissance ! Au temps de ma thèse de doctorat il y a trente ans, mon directeur de recherche, Jean-Claude Latombe, avait réussi à faire organiser par le ministère de la Recherche d'alors, un programme de recherches concertées (PRC) sur l'intelligence artificielle, doté d'un budget d'un million de francs, ce qui représentait alors 20 % de l'effort de recherche du gouvernement en informatique. Depuis, l'intérêt gouvernemental et des organismes publics pour l'IA n'a fait que baisser, l'IA est souvent bien ignorée, ou plutôt enfouie tellement profondément qu'elle n'est plus visible.

Du point de vue de l'enseignement, on a également trop souvent considéré ces dernières années que les outils et techniques de programmation en intelligence artificielle faisaient partie des acquis en algorithmique. Dans certaines villes, l'intelligence artificielle est tellement banalisée que l'enseignement a disparu des formations universitaires, ce qui est grave et le sera encore plus dans le futur si rien n'est fait.

Il faut dire que régulièrement, on nous demande des mots et des challenges nouveaux. On nous dit qu'en intelligence artificielle, on fait toujours un peu la même chose, mais le challenge de l'intelligence artificielle, à savoir l'imitation du comportement humain, est toujours le même et il est bel et bien là, vivace comme aux premiers temps. Alors, en dépit de la réduction des moyens, la communauté résiste néanmoins au risque d'éclatement.

Les sous-communautés de l'IA se sont créées voilà une trentaine d'années parce que le problème d'imitation de l'intelligence humaine était un peu trop compliqué. Mais il faut réussir à faire interagir ces communautés et à intégrer les résultats de leurs recherches et de leurs avancées. Face au risque d'éclatement de la communauté en IA et au besoin d'assurer des rencontres entre les chercheurs, une association, dont l'objet est de « promouvoir et de favoriser le développement de l'intelligence artificielle en France », a été créée en 1993, à l'occasion d'une conférence IJCAI à Chambéry. En 2016, nous comptons jusqu'à 359 membres, dont encore bien trop peu d'industriels. Nous avons élaboré un plan 2011-2019, ayant pour objectifs de désenfouir l'intelligence artificielle et de réunir les communautés spécialisées et d'interagir avec les autres communautés.

Pour ce qui concerne le fait de désenfouir l'intelligence artificielle : nous tenons chaque année une conférence nationale, et organisons deux journées, dont l'une, en octobre, s'intitule « Perspectives et défis de l'intelligence artificielle (PDIA) ». En 2015, le thème en était « Les apprentissages ». En 2016, les « impacts sociaux de l'intelligence artificielle ». L'autre journée que nous tenons, tous les mois d'avril, est un « Forum industriel de l'intelligence artificielle (FIIA) ». L'année dernière, il s'agissait

surtout du lancement du collège industriel de l'AFIA, où les membres ont discuté de l'investissement de l'intelligence artificielle dans leur environnement. Nous publions des bulletins trimestriels, depuis 2011, et avons une page web et des réseaux sociaux. Les dossiers sont accessibles en libre-service sur le *web*, avec des cartographies (agronomie, innovation, robotique, éthique, mégadonnées, réalité virtuelle, médecine, EIH, jeux vidéo, aide à la décision, etc.). Nous commençons aussi l'établissement d'un fonds de fiches car nous sommes très sollicités sur ce sujet et allons faire des efforts pour améliorer la compréhension de l'intelligence artificielle et de ses enjeux au plus grand nombre.

S'agissant de la réunion des communautés spécialisées, une plateforme intelligence artificielle (PFIA) se tient chaque année en juillet. Ayant identifié une dizaine de communautés spécialisées, qui pourront donner lieu à autant de collèges thématiques, cette plateforme permet de se retrouver. En 2017, la PFIA aura lieu à Caen du 3 au 7 juillet, vous y êtes cordialement invités. Deux collèges thématiques sont actifs à ce jour : l'un sur les « Sciences de l'ingénierie des connaissances » et l'autre sur les « Systèmes multi-agents et les agents autonomes ». Nous espérons lancer trois nouveaux collèges cette année. Par ailleurs, notre collège industriel se charge du forum industriel de l'IA ainsi que de la conférence sur les applications pratiques de l'IA au sein de la PFIA. Depuis deux ans, nous réalisons aussi des compétitions d'intégration au sein de la PFIA, qui constituent des mécanismes incitatifs pour que les collègues des dix collèges interagissent autour d'un objectif commun, et intègrent leurs technologies respectives. Pour l'instant, il s'agit de compétitions académiques. En 2017, la compétition porte sur l'intelligence artificielle dans les jeux interactifs. Mais nous espérons que ces compétitions, à l'avenir, aient par exemple pour objectif de résoudre des problèmes de nature industrielle qui seraient soumis par l'un des membres de notre collège industriel.

Enfin, nous cherchons à interagir avec les autres communautés, *via* des journées communes avec d'autres sociétés savantes et associations, d'organisations patronales (exemple de la journée commune avec le MEDEF le 23 janvier 2017), des groupements de recherche (GdR) du CNRS, mais aussi des pôles de compétitivité. Nous avons planifié une journée avec la société de philosophie des sciences (SPS) le 2 février, une journée avec l'AFIHM sur l'interaction homme-machine le 17 mars, une journée EIAH et IA avec l'ATIEF en juin, une journée « jeux informatisés » avec le pré-GdR AFAIA du CNRS en juillet, et, enfin, une journée éthique avec le COMETS du CNRS et la CNIL puis une journée recherche opérationnelle avec la ROADEF en septembre. Nous organisons également le 27 avril notre forum FIIA et le 6 octobre notre journée PDIA. Enfin, nous sommes consultés régulièrement, par exemple par l'Académie des technologies, par différents instituts comme l'IHEDN, par des municipalités et associations qui s'inquiètent de l'impact de l'intelligence artificielle, et par la presse, qui relaie les progrès en intelligence artificielle, avec un accent sur les États-Unis.

Pour conclure et envisager l'avenir de l'intelligence artificielle, nous pensons que cette dernière restera une force tranquille durant les soixante prochaines années car nous sommes encore loin d'imiter le comportement humain. Ceci dit, nous estimons que les machines sont, d'année en année, plus intelligentes, alors que l'être humain, lui, évolue moins vite. Je suis très surpris de voir que l'on se réfère toujours aux États-Unis, alors qu'il est nécessaire de considérer ce qui se passe en Asie, en particulier en Chine et au Japon.

Aux États-Unis, l'industrie automobile investit beaucoup en intelligence artificielle et l'administration de Barack Obama a fait un rapport pour se préparer à l'intelligence artificielle, mais les gouvernements japonais et chinois ne sont pas en reste. C'est en Chine que l'investissement public semble le plus conséquent. En mai 2016, le nouveau plan national a ainsi intégré un programme en IA sur trois ans. J'ai été personnellement fasciné par le lancement du drone Ehang 184, en janvier 2016. Ce drone pèse 200 kilos et peut transporter une personne qui indique *via* un écran tactile la destination à laquelle il souhaite se rendre avec une autonomie de 25 minutes. Les Chinois ont beaucoup investi, et ont une vision différente, et à certains égards complémentaire, de celle des Américains. Le Japon a, lui aussi, lancé un nouveau plan d'action en intelligence artificielle sur trois ans. La robotique est dominante dans leurs investissements. Il y a plusieurs raisons à cela parmi lesquelles le fait fascinant que le shintoïsme conduit à croire dans la réincarnation dans des objets. La relation d'un Japonais avec un robot est donc assimilable à celle qu'il peut avoir avec un autre être humain. L'acceptabilité de la robotique au Japon ne sera sans doute jamais atteinte en France. Il faut accepter ces différences et cultiver nos spécificités. Nous avons une vision française de l'intelligence artificielle à valoriser nous aussi.

Au minimum, nous espérons que la France rétablira le poids de l'intelligence artificielle à 20 % au moins de l'informatique, ce qui est bien loin d'être le cas actuellement. La complexité de toutes les thématiques qui la composent est telle qu'un soutien de l'État à la hauteur des ambitions est nécessaire. Il incombe aux responsables politiques de fournir les moyens nécessaires à une intelligence artificielle française digne de ce nom. Il me semble, par ailleurs, essentiel de rétablir l'enseignement de l'intelligence artificielle dans toutes les filières universitaires. 40 % des chercheurs en apprentissage profond avouent ne pas savoir expliquer les raisons pour lesquelles leur technologie donne des résultats. Les machines vont devenir de plus en plus complexes et de moins en moins explicables. Il y a donc là un vrai problème, qui se posera de plus en plus demain.

Il convient aussi de mettre en valeur les réalisations françaises. Environ 500 personnes se sont inscrites à la journée « Entreprises de France et intelligence artificielle » que nous organisons la semaine prochaine avec le MEDEF. Des entreprises françaises y présenteront leurs réalisations en intelligence artificielle. Il faudrait que nos entreprises communiquent aussi

bien que le font les GAFAs sur leurs grandes réussites, au lieu de râler sur le coût du travail.

Enfin, il importe d'établir une meilleure coordination entre tous les acteurs de l'intelligence artificielle car si on perd la main au niveau national, il y a clairement un risque de perte de souveraineté. Chaque année, la recherche avance bien, avec différents cycles : on a beaucoup parlé à un moment donné de systèmes experts ; actuellement on parle de *deep learning*.

Les États-Unis ont compris leur intérêt stratégique à l'intelligence artificielle, et ils imposent leurs vues. C'est très visible dans les conférences internationales, en particulier, ainsi que lors des auditions.

Suivant les acceptions, l'intelligence artificielle peut représenter 50 % de l'informatique. La moitié des *Lecture Notes in Computer Science* aux éditions Springer-Verlag est consacrée aux *Lecture Notes in Artificial Intelligence*. En France, on n'en est pas là. Lorsque je faisais ma thèse, mon directeur de recherche, Jean-Claude Latombe, avait réussi à organiser voilà trente ans un PRC-GDR (programme de recherches coordonnées - groupe de recherche), avec un budget d'un million de francs, ce qui représentait 20 % de l'effort de recherche en informatique.

Dans certaines villes, l'intelligence artificielle est tellement banalisée que l'enseignement a disparu des formations universitaires, ce qui est grave. Depuis quelques années, on nous demande des mots nouveaux, des nouveaux challenges. En réalité, le challenge de limitation du comportement humain est encore présent. On nous oppose le constat que l'on fait toujours un peu la même chose et que la technologie est quelque peu enfouie dans le système intelligent. Même si cette situation a éclaté la communauté, celle-ci résiste néanmoins.

Les thèmes de l'intelligence artificielle au niveau international concernent non seulement le *deep learning*, mais également la planification, la modélisation, les connaissances, etc. Devant l'éclatement de la communauté et, simultanément, à la nécessité que les chercheurs en intelligence artificielle se rencontrent, a été créée en 1993, à l'occasion de la tenue de la conférence IJCAI (*International Joint Conference on Artificial Intelligence*) à Chambéry, une association régie par la loi de 1901, dont l'objet est la promotion et le développement de l'intelligence artificielle en France.

Nous avons élaboré un plan 2011-2019, ayant pour objectifs de désenfouir l'intelligence artificielle et de réunir les communautés spécialisées. Des sous-communautés se sont créées voilà une trentaine d'années parce que le problème était un peu trop compliqué. Mais il faudrait réussir à intégrer les résultats des recherches de toutes ces communautés. D'ailleurs, l'intelligence artificielle n'est pas un monobloc. Même si ce domaine fait partie d'une grande part de l'informatique, il doit y avoir une interaction.

En France, l'AFIA a identifié dix thèmes. Certes, la vision peut paraître un peu académique, mais 95 % des 348 membres de l'association sont des académiques, contre très peu d'industriels. L'apprentissage profond est, à nos yeux, une partie de l'apprentissage – moins de 50 %. Nous mettons actuellement un focus particulier sur l'apprentissage profond, mais cela pourrait tout aussi bien concerner le véhicule autonome, dont les recherches accusent un peu de retard. Certes, on peut toujours imaginer les prochains challenges, mais, dans quelques années, la priorité sera donnée à un autre sujet.

Non seulement l'informatique est exposé entre plusieurs dizaines d'associations, mais l'intelligence artificielle déborde sur un certain nombre de domaines. Nous interagissons donc avec d'autres associations.

Pour désenfourer l'intelligence artificielle, nous organisons une conférence nationale, des journées propres, des bulletins trimestriels incluant des dossiers. Nous commençons aussi l'établissement d'un fonds de fiches. Nous sommes en effet très sollicités sur ce sujet ; nous allons donc faire des efforts en termes de communication pour simplifier les choses. Nous avons aussi un portail *web* et des réseaux sociaux. J'indique que les dossiers sont accessibles en libre-service sur le *web*, avec des cartographies. Des équipes travaillent sur des thématiques applicatives données.

Réunir les communautés spécialisées : on est en train de créer autant de collèges thématiques qu'il y a de thèmes, mais on a aussi créé, il y a un an, un collège industriel. À cet égard, nous organisons lundi prochain une journée commune avec le MEDEF.

Par ailleurs, entre quatre cents et quatre cent cinquante personnes sont attendues l'été prochain pour assister à la plateforme intelligence artificielle. De plus, des journées communes sont organisées avec d'autres associations, des organismes, des pôles de compétitivité, des organisations patronales, des académies, des organisations gouvernementales, les municipalités, des organes de presse écrite et orale.

Une journée commune avec la société de philosophie des sciences est aussi prévue en février. Et bien d'autres rencontres sont programmées, avec d'autres associations, des groupements de recherche du CNRS ou tout organisme qui le demanderait.

Quoi qu'il arrive, on espère que l'intelligence artificielle continuera à être une force tranquille durant les soixante prochaines années. On est très loin d'être arrivé à étudier la limitation du comportement humain. On le constate, les machines sont chaque année de plus en plus intelligentes, alors que l'être humain ne va pas beaucoup évoluer. Il est nécessaire de considérer ce qui se passe en Asie. À cet égard, je suis très surpris de voir que l'on se réfère toujours aux GAFAs, notamment. Pour ma part, je regarde ce qui se passe en Chine et au Japon. Vous le savez, le gouvernement Obama a fait un rapport de prospective pour se préparer à l'intelligence artificielle. Mais, en mai 2016, le gouvernement japonais a lancé un grand plan dans ce domaine.

La Chine a également engagé en juin dernier un programme national en la matière pour les trois prochaines années.

J'espère que la France rétablira le poids de l'intelligence artificielle à 20 % au moins de l'informatique, ce qui est bien loin d'être le cas actuellement. La complexité de toutes les thématiques est telle que si le soutien de l'État n'est pas à la hauteur des ambitions, il sera alors quasiment sans effet. À vous de voir si vous y mettez les moyens.

Il me semble essentiel de rétablir l'enseignement dans les formations universitaires. Certes, on parle beaucoup actuellement de l'apprentissage profond. Mais sachez que 40 % des chercheurs en la matière ne savent pas expliquer les raisons pour lesquelles cela marche. Les machines vont devenir encore plus intelligentes et si aucune formation n'est dispensée en intelligence artificielle, il arrivera un moment où l'on comprendra encore moins ce qui se passe. Se pose donc là un problème.

Il convient de mettre en valeur les réalisations françaises. Je suis très heureux de voir que cinq cent dix personnes se sont inscrites à la journée que nous organisons la semaine prochaine en commun avec le MEDEF. On montrera aux entreprises françaises les réalisations en intelligence artificielle. Cela me semble important. Il faudrait que celles-ci communiquent aussi bien que le font les GAFAs sur leurs grandes réussites, au lieu de râler.

Ensuite, il importe d'établir une meilleure coordination entre les acteurs de l'intelligence artificielle - on peut toujours le souhaiter. Je vous l'ai dit, si l'on perd la main au niveau national ou au niveau scientifique, il y a clairement un risque de perte de souveraineté. Il faut sans doute considérer l'intelligence artificielle comme une technologie de souveraineté.

M. Claude de Ganay, rapporteur. - Monsieur Bertrand Braunschweig, vous êtes directeur du centre de l'Institut national de recherche en informatique et en automatique (Inria) de Saclay et représentez le président-directeur général d'Inria, Antoine Petit, aujourd'hui retenu. Inria a publié l'été dernier un livre blanc sur l'intelligence artificielle, qui fait le point sur l'état de la recherche à ce sujet au sein de votre institut.

6. M. Bertrand Braunschweig, directeur du centre Inria de Saclay

J'ai effectivement coordonné la réalisation de ce premier livre blanc sur l'intelligence artificielle, qui présente les équipes d'Inria actives dans les différents domaines de l'intelligence artificielle. L'Institut est organisé en équipes-projet, au nombre de deux cents environ, sur nos huit centres de recherche, et qui regroupent chacune une vingtaine de chercheurs en moyenne. Un tiers de ces équipes touche, de près ou de loin, au domaine de l'intelligence artificielle, une proportion qui a significativement augmenté

ces dernières années, avec la montée en puissance de l'apprentissage-machine. Plus d'un millier de communications en conférence et plus de quatre cents articles dans les revues consacrées à l'intelligence artificielle ont été publiés par nos chercheurs au cours des dix dernières années.

Ce livre blanc mentionne quelques faits marquants qui motivent l'intérêt actuel pour l'intelligence artificielle, et rappelle aussi qu'il y a des questions de société à traiter. Inria a pris des initiatives en ce sens depuis plusieurs années, en créant, au sein de l'alliance Allistene, la commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique (CERNA), dont le premier rapport public traite de la recherche en robotique, ainsi qu'en se dotant d'un organe interne, le comité opérationnel d'évaluation des risques légaux et éthiques, qui traite les cas pratiques soumis par nos équipes de recherche. En cette année où nous célébrons les cinquante ans d'Inria, nous entendons contribuer au débat public qui doit exister sur l'intelligence artificielle.

L'année 2017 est aussi celle de la préparation de notre plan stratégique scientifique pour la période 2018-2022. Ce plan sera organisé autour de défis scientifiques motivants, dont plusieurs porteront, soyez-en certains, sur l'intelligence artificielle. Nous finissons, en ce moment, un nouveau livre blanc sur le véhicule autonome et connecté : un des défis du plan devrait porter sur ce sujet.

Le livre blanc est structuré en huit sous-domaines, sur lesquels nos équipes sont actives et qui couvrent une bonne partie de la recherche en intelligence artificielle. On y trouve, entre autres entrées, les défis génériques de l'intelligence artificielle, l'apprentissage automatique, l'analyse des signaux - vision et parole, notamment -, les connaissances et le *web* sémantique, la robotique et les véhicules autonomes, les neurosciences et la cognitive, le traitement du langage, la programmation par contrainte pour l'aide à la décision.

Je m'en tiendrai à formuler quelques remarques sur certains secteurs qui me paraissent particulièrement importants. En matière d'apprentissage automatique, la révolution de l'intelligence artificielle est essentiellement venue de progrès scientifiques considérables en matière de traitement du langage et de la parole, de vision, de robotique et, surtout, d'apprentissage automatique. En cette dernière matière, même si des résultats remarquables ont été obtenus, il reste encore de nombreux défis propres à motiver longtemps les chercheurs.

Je pense, en particulier, à l'apprentissage non supervisé, sans intervention d'un oracle humain, à l'apprentissage sous contrainte, par exemple pour la gestion de la vie privée, à l'apprentissage de causalité, indispensable pour construire des systèmes prédictifs, ou encore à l'apprentissage continu et sans fin, pour des systèmes destinés à opérer vingt-quatre heures sur vingt-quatre, sept jours sur sept.

Autre défi, l'interaction avec les humains.

Les systèmes d'intelligence artificielle étant appelés à interagir avec des utilisateurs humains, ils doivent être capables d'expliquer leur comportement, de justifier les décisions qu'ils prennent, faute de quoi, ces systèmes ne seront pas acceptés, par manque de confiance. J'ajoute que les systèmes d'intelligence artificielle ont besoin d'une certaine flexibilité et doivent être capables de s'adapter à différents utilisateurs et différentes attentes. Il importe donc de développer des mécanismes d'interaction favorisant une bonne communication entre humains et systèmes d'intelligence artificielle.

J'en viens au défi de l'ouverture à d'autres disciplines. Sachant que l'intelligence artificielle sera souvent intégrée dans un système comportant d'autres éléments, les spécialistes du domaine devront collaborer avec ceux d'autres sciences de l'informatique – modélisation, vérification, validation, visualisation, interaction machine –, ainsi que d'autres disciplines comme la psychologie, la biologie, les mathématiques, mais aussi les sciences humaines et sociales – économie, ergonomie, droit. Le mieux serait que les chercheurs possèdent une double, voire une triple compétence, afin de faire sauter les barrières disciplinaires, qui rendent malheureusement difficile la réalisation de systèmes complets. J'ai bon espoir que les instituts Convergence, financés par les investissements d'avenir, y contribuent. En attendant, il est indispensable de composer des équipes pluridisciplinaires pour aborder les problèmes sous tous les angles.

Composante incontournable des systèmes critiques, la certification des systèmes d'intelligence artificielle ou leur validation par des moyens appropriés constitue également un véritable défi. Alors que la vérification, la certification et la validation des systèmes classiques sont déjà des tâches difficiles, même s'il existe des outils exploitables, l'application de ces outils aux systèmes d'intelligence artificielle comportant notamment des systèmes d'apprentissage automatique est une tâche encore plus ardue, à laquelle il convient de s'attaquer si nous voulons utiliser ces systèmes dans des environnements tels que ceux de l'avion, des centrales nucléaires, des hôpitaux...

J'en arrive à l'interaction entre l'apprentissage et la modélisation, sujet qui nous tient particulièrement à cœur. Alors que les scientifiques passent beaucoup de temps à concevoir des modèles – physiques, mathématiques, symboliques –, les algorithmes d'apprentissage ignorent encore assez largement ces modèles, et travaillent sur des données brutes. Il faut, à notre sens, briser cette barrière entre apprentissage et modélisation, pour établir une interaction dans les deux sens, afin d'améliorer les performances de l'apprentissage automatique grâce aux connaissances synthétisées dans des modèles, ce qui permettra, en retour, d'améliorer ou de spécialiser ces modèles, grâce aux résultats de l'apprentissage. Cela peut requérir une interface entre les couches numérique et symbolique : la tâche n'est pas facile, mais c'est un peu le Graal dont la poursuite doit nous occuper.

Parmi les autres défis importants, sur lesquels je pourrais revenir au cours du débat, il faut également compter la gestion des données multimodales provenant de différents capteurs, la compréhension de scènes et d'environnements, en particulier pour les robots et les véhicules autonomes, et, toujours pour les robots, la prise en compte de l'incertitude et des données incomplètes ou disponibles seulement à une certaine fréquence. Je pense également, pour le domaine du *web* sémantique, à la connexion entre les ontologies, qui en constituent le moteur, et les données stockées dans les bases, qui en sont le carburant.

Le plus grand défi, pour conclure, n'est non pas technique, mais stratégique. La France excelle sur tous ces sujets en raison de la qualité de ses scientifiques. Tant pour les approches mathématiques - statistique, modélisation, optimisation, apprentissage - que pour la modélisation et la représentation des connaissances, nous bénéficions d'une longue tradition de raisonnement logique et de cartésianisme. Pour preuve, les chercheurs français suscitent de plus en plus la convoitise des entreprises internationales, qui offrent des niveaux de salaire et de financement extrêmement motivants. Mettre nos capacités au service du développement économique grâce au transfert technologique doit permettre à nos entreprises de prendre des parts d'un marché international en forte croissance. Nos entreprises, grandes et petites, ont besoin de ces technologies ; nos chercheurs en intelligence artificielle sont parmi les meilleurs au monde : il faut tisser plus de liens entre ces deux mondes. Cela pourrait être la mission principale d'un grand plan national en intelligence artificielle. Demain sera d'ailleurs lancée la préparation de la stratégie nationale en intelligence artificielle, sous le patronage des deux ministères concernés, celui de l'industrie et celui de la recherche. Inria estime que notre pays doit se montrer ambitieux et faire de l'intelligence artificielle une vraie priorité. Au vu des investissements que consentent plusieurs grands pays, c'est au moins un milliard d'euros sur dix ans qu'il faut mobiliser, aides publiques et contribution des acteurs économiques confondues ; je pense aux entreprises, mais aussi aux collectivités territoriales, qui pourraient fournir des sites d'expérimentation grandeur réelle. Tel est le principal défi qu'Inria vous propose de relever.

M. Claude de Ganay, rapporteur. - Monsieur David Sadek, vous êtes directeur de la recherche de l'Institut Mines-Télécom. Enseignant-chercheur et spécialiste en intelligence artificielle, vous avez mené une carrière atypique, puisque vous êtes d'abord entré chez Orange, dont vous êtes devenu, il y a dix ans, le directeur délégué à la recherche avant de revenir à la recherche publique.

7. M. David Sadek, directeur de la recherche à l'Institut Mines-Télécom

Permettez-moi d'abord d'évoquer les publications de l'Institut Mines-Télécom, notamment son cahier de veille sur l'intelligence artificielle, qui offre une sorte de patchwork de points de vue sur le sujet. Créé conjointement avec la Fondation Télécom, ce cahier témoigne de la diversité de nos activités dans le domaine de l'IA, qui touchent au traitement des données massives – en la matière, plusieurs chaires industrielles sont portées par l'Institut –, à l'apprentissage, à la reconnaissance des formes, à la compréhension de scènes, à la représentation des connaissances, aux neurosciences informationnelles – domaine nouveau qui vise à caractériser l'information mentale et corticale –, à la robotique, aux réflexions éthiques.

Si j'aborde dans mon intervention la question de l'apprentissage, sur laquelle les orateurs m'ayant précédé ont également mis l'accent, je n'en estime pas moins qu'il faille se garder de réduire l'intelligence artificielle à un seul de ses pans.

L'intelligence artificielle est un vaste domaine, pluridisciplinaire, dont l'unité tient en ceci que l'on y cherche à inculquer à des machines des compétences cognitives qui sont le propre de l'humain. Comme je le dis souvent, l'intelligence artificielle, c'est la didactique des machines : il s'agit de parvenir à expliquer à une machine comment faire des choses que l'humain fait très bien. Cela pose un ensemble de défis, au nombre de six ou sept.

Bertrand Braunschweig a évoqué des défis génériques ; je me situerai à un échelon plus « primitif ». Gérard Sabah a parlé de la représentation du sens : ce n'est pas parce qu'une machine parvient à classer une image de chat dans un groupe d'images similaires qu'elle a compris ce qu'était un chat. Certes, on avance dans le traitement sémantique de l'information, et d'autres approches essaient de tendre vers une représentation du sens au moyen de ce que l'on appelle des ontologies ou des réseaux sémantiques, mais il n'en demeure pas moins que la représentation du sens reste encore un défi pour l'intelligence artificielle.

Autre défi, la notion de sens commun.

L'être humain se caractérise par sa capacité à raisonner sur la base d'informations approximatives, en menant des raisonnements symboliques du type « en général » : « En général il fait beau en Bretagne », assertion au reste discutable... Il use de raisonnements par défaut, mais aussi de raisonnements par analogie : « Paris est à la France ce que Washington est aux États-Unis », par exemple. Mais, dans le domaine de l'intelligence artificielle, de tels modes de raisonnement ne sont pas encore assimilés, y compris par les approches que l'on met aujourd'hui en avant.

Autre gageure, évoquée par Jean-Gabriel Ganascia, le sujet des agents conversationnels : l'interaction, le dialogue avec la machine. C'est en

soi un concentré de défis. Il n'est d'ailleurs pas anodin que le fameux test de Turing fasse appel à du dialogue humain-machine. Sont en jeu des comportements qui supposent une compréhension du langage naturel, une interprétation en contexte, qui fait sortir du simple jeu des questions-réponses, etc. N'allez pas croire qu'un logiciel comme Siri a résolu la question ; on en est loin. Avoir avec une machine un dialogue évolué reste encore un défi pour l'intelligence artificielle, quelle qu'en soit l'approche. Outre les compétences que je viens de citer, il faut aussi y ajouter des capacités de raisonnement, d'introspection, de production de langage, etc. Gérard Sabah a rappelé que l'on est loin encore de comprendre ce qu'est la notion de conscience – elle reste d'ailleurs, dans la sphère de l'humain, indécidable pour tout autre que soi – ; ce qui doit nous importer, c'est plutôt le rôle causal des aptitudes cognitives attribuées à une machine.

Sans aller donc jusqu'à traiter la question de la conscience, celle de l'introspection, c'est-à-dire la capacité, pour une machine, à raisonner et à réfléchir sur ses propres connaissances et ses propres comportements, reste un défi pour l'intelligence artificielle.

Vient ensuite ce que l'on appelle l'intelligence émotionnelle, à savoir la capacité à caractériser et à reconnaître des émotions.

À quoi il convient encore d'ajouter la capacité d'expliquer les comportements, d'en rendre compte. C'est là un point très important car un système d'intelligence artificielle, pour être acceptable, doit être capable d'expliquer ce qu'il est en train de faire et pourquoi. Une machine, autrement dit, doit être au fait de ce qui régit son fonctionnement.

Autre point important, ce que l'on appelle la « prouvabilité » des systèmes, soit la capacité à montrer qu'un système intelligent fait ce que l'on attend de lui, et seulement ce que l'on en attend. Alors que l'on se soucie beaucoup, aujourd'hui, des règles éthiques, il serait malvenu que des systèmes intelligents soient utilisés à des fins autres que celles pour lesquelles ils ont été initialement conçus. Or des comportements contraires à l'éthique peuvent émerger de manière imprévue dans un système dont les propriétés n'auraient pas été prouvées, au sens que je viens de définir, *ex ante*. Je cite souvent l'exemple d'un système de bataille navale utilisé, du temps où j'étais jeune chercheur, par le Département de Défense américain, pour entraîner des pilotes de flotte navale. On s'est aperçu que, si le système gagnait tout le temps, c'est parce qu'il détruisait les bateaux touchés de sa propre flotte, pour qu'ils ne la ralentissent pas ! On avait pourtant simplement indiqué au système qu'il fallait maximiser la vitesse de la flotte, et il a fait de lui-même cette inférence simple qu'un bateau touché la ralentissait. Voilà ce qui peut advenir lorsque la validité d'un système intelligent n'a pas été prouvée. La recherche de l'objectif de maximisation peut conduire à des actions non éthiques.

J'insiste, pour conclure, sur l'hybridation des approches, de nature à faire avancer la recherche. Les approches, qu'elles soient neuro-inspirées,

connexionnistes, stochastiques ou, au contraire, cognitivistes ou symbolistes ont toutes à apporter dans la compréhension, la modélisation et la mise en œuvre des comportements intelligents. Nous avons des équipes excellentes, en France, dans la plupart de ces pans de l'intelligence artificielle, et nous devrions nous investir pour que ces équipes travaillent ensemble. Nous sommes à l'aube d'une ère nouvelle. Cela exige de faire la part des choses, de trier entre le bon grain et l'ivraie. Les approches par apprentissage, notamment le *deep learning*, sont très performantes pour tout ce qui relève de la perception – reconnaissance de sons, de parole, d'images, etc., traitement de signaux pour la reconnaissance de formes –, tandis que ce qui a trait à la mobilisation de comportements intelligents, relevant plus de la cognition, tels que la conduite de dialogues humain-machine, et qui passe par des notions d'attitudes mentales et de raisonnement symbolique, reste largement à défricher par ces approches : la combinaison de méthodes d'inspiration cognitive et d'approches faisant appel à la classification stochastiques serait, en la matière, fructueuse.

M. Claude de Ganay, rapporteur. – Monsieur Jean-Daniel Kant, vous êtes maître de conférences à l'Université Pierre-et-Marie-Curie Paris-VI. Spécialiste de la technologie, peu connue, des systèmes multi-agents, vous vous intéressez à l'hybridation entre technologies d'intelligence artificielle.

8. M. Jean-Daniel Kant, maître de conférences à l'Université Pierre-et-Marie-Curie Paris-VI

Je vous remercie de votre invitation et de l'initiative de l'OPECST. C'est une excellente idée de nouer le dialogue entre chercheurs et parlementaires, et j'espère que cette occasion qui nous est donnée de nous rencontrer sera suivie de bien d'autres.

Je vais m'efforcer de vous exposer ce que sont les systèmes multi-agents. Le *deep learning*, le *machine learning*, les robots sont la face visible de l'intelligence artificielle, qui comporte cependant d'autres aspects, évoqués par David Sadek et Bertrand Braunschweig. L'intelligence artificielle, de fait, se décline au pluriel. L'une de ses approches se situe du côté des données, tandis que l'autre se tourne plutôt vers les comportements. La première approche vise, grâce au traitement d'un très grand nombre de données, à faire émerger des comportements liés à l'intelligence humaine. Elle est beaucoup utilisée pour la reconnaissance de formes, mais aussi en aide à la décision. Les banques commencent ainsi à utiliser de tels systèmes pour décider à quels clients accorder un prêt. Les assurances s'y lancent également, pour détecter les fraudes.

Les réseaux de neurones – l'appellation est métaphorique – sont très puissants, mais ce sont des boîtes noires. Les spécialistes, comme Yann

LeCun, disent, honnêtement, qu'ils ne savent pas vraiment comment cela fonctionne. Il y a là une question politique, une question citoyenne. Qui voit son banquier rejeter sa demande de prêt, s'il s'entend répondre, quand il demande les raisons du refus, que c'est parce que le neurone n° 39 et la connexion 48 étaient inférieurs à 0,4, est en droit de s'offusquer, et pourrait bien se pourvoir en justice.

Vous le savez, l'Union européenne a été réactive en ce domaine, puisqu'un règlement et une directive adoptés récemment concernent l'explication des décisions automatisées. Bientôt, tous les systèmes devront être capables de justifier leur comportement. Cela sera sans nul doute un frein à l'utilisation, dans l'aide à la décision, des systèmes à réseaux de neurones.

L'autre approche, qui n'est pas gouvernée par les données, passe par la modélisation des comportements humains. Elle s'efforce, dans une démarche pluridisciplinaire, faisant appel à des psychologues, des économistes, des sociologues, de comprendre les mécanismes qui régissent les comportements humains, pour tenter de les modéliser dans des programmes informatiques.

Un système multi-agents est fait de programmes autonomes, dotés de certaines capacités cognitives, qui peuvent être très simples ou très complexes. Ces programmes, que l'on appelle « agents », communiquent entre eux, interagissent, pour former une sorte d'intelligence collective. Cela est très utile lorsque l'on veut modéliser des systèmes humains, sociétés ou économies.

L'autre question qui vaut d'être posée est celle des objectifs de l'intelligence artificielle. J'en vois deux. L'un est de faire des machines intelligentes, autonomes, destinées, même si on ne le présente pas ainsi, à remplacer les êtres humains. Quand on crée une voiture autonome, plus besoin de chauffeur. Dans tous les cas où l'on considère que l'être humain sera moins performant que la machine, ou dans les cas où l'on ne peut employer qu'un robot – pour une expédition vers Mars, par exemple –, cela fait sens. Mais la société, le monde politique doivent s'interroger : est-ce le cas dans tous les domaines ? D'autant que remplacer les humains par des machines n'est pas sans incidence sur le marché de l'emploi, c'est mettre des gens au chômage, et c'est bien pourquoi Barack Obama évoquait l'idée de revenu universel.

Le second objectif de l'intelligence artificielle est non pas de remplacer l'homme, mais de lui apporter une assistance. Quand on dialogue avec Siri, on obtient de l'aide. En viendra-t-on un jour à ne plus dialoguer qu'avec des machines ? C'est une autre histoire.

Quoi qu'il en soit, mes recherches sont plutôt orientées vers cet objectif d'aide à la décision. Quelles en sont les applications potentielles ? Elles peuvent concerner, par exemple, l'économie. C'est ainsi qu'avec l'économiste Gérard Ballot nous avons développé un simulateur multi-

agents du marché du travail français. Grâce à une modélisation à 1/2 000^e de ce marché, nous sommes en mesure d'évaluer des politiques publiques. Nous avons, par exemple, évalué la loi El Khomri, et les résultats, assez surprenants, font apparaître que les choses sont beaucoup plus compliquées que ce qui se disait alors. Cette évaluation a donné lieu à bien des réactions sur le *web*, les gens se demandant pourquoi les ministères n'utilisaient pas de tels systèmes pour faire leurs études d'impact. Certes, être totalement prédictif sur une matière aussi complexe que le marché du travail n'est pas facile, mais il reste que les systèmes multi-agents sont une voie.

Nous avons également créé des applications dans le domaine social pour étudier la diffusion de l'innovation dans une population de consommateurs, ou bien la dynamique des opinions et des attitudes – nous réfléchissons actuellement à un projet destiné à étudier la radicalisation sur Internet.

Créer des systèmes fonctionnant comme un miroir de l'activité humaine est un moyen, à mon sens, pertinent pour rendre compte de leur complexité. On peut également hybrider, et étudier les interactions entre l'humain et ces technologies.

C'est une voie prometteuse, mais qui demande des moyens. En France, hélas !, nous en manquons. Ce serait pourtant le moyen de garantir notre souveraineté, par rapport à des pays comme les États-Unis ou la Chine. J'insiste aussi sur l'utilité de la recherche pluridisciplinaire, qui n'est pas assez défendue.

M. Claude de Ganay, rapporteur. – Monsieur Benoît Le Blanc, vous êtes directeur adjoint de l'École nationale supérieure de cognitive. Vous êtes spécialiste en intelligence artificielle et assumez actuellement une mission pour le ministère de l'enseignement supérieur et de la recherche, dont vous nous direz sans doute quelques mots.

9. M. Benoît Le Blanc, directeur adjoint de l'École nationale supérieure de cognitive

Beaucoup a déjà été dit sur les technologies d'intelligence artificielle, un sujet dont on débat dans les médias. La maturité de ces technologies, issues de laboratoires de recherche, est-elle suffisante pour que des entreprises engagent des moyens destinés à développer massivement des produits qui vont bouleverser notre quotidien ? Telle est la question.

À titre personnel, je dois dire que je suis impressionné par le programme Watson et les capacités de développement informatique que met à disposition IBM, au travers de son logiciel *Bluemix*. Un tel logiciel permet

de développer des applications de « *cognitive computing* », pour, par exemple, transformer le langage en texte, ou inversement.

À lire rapidement les journaux et sites *web* d'actualité, on pourrait considérer qu'il existe, dans le paysage de l'intelligence artificielle, deux groupes d'acteurs, l'un constitué par les chercheurs « à l'ancienne » que nous sommes, l'autre par de jeunes découvreurs de la French Tech, qui lancent des start-up sur tous les sujets, en y injectant de l'intelligence artificielle.

Or les technologies d'intelligence artificielle sont pour certaines destinées à des usages par le public et, dès lors qu'intervient le facteur humain, les choses sont beaucoup plus complexes qu'il n'y paraît. De fait, les programmes d'intelligence artificielle sont destinés à converser avec les personnes. Nous avons, en France, beaucoup de talents en sciences humaines et sociales, des laboratoires très performants, peuplés de chercheurs très technophiles, qui peuvent nous apporter beaucoup. Cela suppose, et j'insiste, d'être capable de s'affranchir des barrières disciplinaires qui nous enferment dans certains modes de pensée. Nous sommes passés, entre la fin du XX^e siècle et l'aube du XXI^e siècle, des sciences de l'information aux sciences de la connaissance, lesquelles supposent d'être capable de donner un contexte et du sens à l'information.

La cognitive est ainsi à l'informatique ce que la connaissance est à l'information. Il s'agit dans la cognitive d'adapter la technologie aux capacités, aux limites, aux préférences humaines, pour en rendre l'usage plus simple. L'École nationale supérieure de cognitive, au sein de Bordeaux INP, développe un programme d'enseignement en ingénierie lié à notre programme de recherche sur la cognition et aux entreprises passionnées par le sujet et hébergées sur notre campus. Nous travaillons sur l'expérience utilisateur, soit la manière dont les gens interagissent avec les technologies et au-delà, nous développons des programmes du futur autour du partage des connaissances et de l'hybridation humain-machine au service de la cognition. Autrement dit, il s'agit de réfléchir, d'une part, à la manière dont les gens vont partager les connaissances entre eux, et, d'autre part, vont intégrer à leur personne même une composante technique.

Si j'observe la manière dont nous avons mené chacun préparé et mené nos exposés ce matin, je constate que certains utilisent un diaporama pour appuyer leur discours, quand d'autres préfèrent s'en remettre à des notes manuscrites, imprimées ou bien lues sur l'écran de leur ordinateur. Nous sommes tous informaticiens et, pourtant, nous restons, mentalement, pluriels. Nos approches de la technologie sont diverses. De la même manière, c'est en mettant au creuset de l'interdisciplinarité sciences de l'information et de la cognition et sciences sociales et humaines que l'on créera des synergies.

Par exemple, nous butons, à l'heure actuelle, sur la capacité à relier réseaux sémantiques et réseaux neuronaux. Cela signifie que les modèles mentaux ancrés chez les chercheurs sont profondément différents. Certains

sont « symbolistes » – j’en fais partie, comme plusieurs autour de cette table – et pensent que l’on crée de l’intelligent en manipulant des symboles, d’autres sont « émergentistes », et travaillent sur les réseaux de neurones en pensant que l’on crée de l’intelligent par association de petites unités en configurations : il faut pouvoir travailler ensemble, même s’il est difficile de comprendre les présupposés et les postures intellectuelles de chacun.

L’intelligence artificielle et la cognition sont des domaines de recherche transdisciplinaires, au même titre que la santé publique, les études sur le climat ou l’alimentation ; tous ces domaines appellent des expertises larges. Mais le problème est que l’on appréhende la transdisciplinarité par le biais de chercheurs attachés à une discipline. Il n’existe pas d’instance qui permette d’évaluer sereinement et positivement les travaux que mènent des chercheurs quand ils changent de section, de bain disciplinaire, si bien que le faire représente un sacrifice de carrière. Il faut parvenir à lever ce frein.

La transdisciplinarité doit nous guider vers des recherches axées sur la performance combinée entre le cerveau humain et le calcul de la machine. Plus que d’une acceptabilité des machines, je parlerai d’une appropriation, qui se fera lorsque les gens tireront bénéfice de l’usage de leur propre machine.

J’ai eu le privilège de discuter avec deux personnalités qui m’ont marqué : Anatoli Karpov, il y a quelques années, lors de l’une de ses conférences, où il expliquait très clairement que, dès lors que les machines battent les humains aux échecs, l’avenir du jeu d’échecs passe par la confrontation de joueurs utilisant des machines, et, plus récemment, Fan Hui, champion d’Europe de jeu de Go qui a alimenté la machine par laquelle le champion du monde, Lee Sedol, a été battu. Fan Hui, appelé à tester cette machine, dit s’être retrouvé face à elle comme face à un mur. La machine l’a battu, et il a, raconte-t-il, éprouvé une terrible honte. Cela a été pour lui d’autant plus difficile à vivre qu’il a dû garder le secret jusqu’au combat contre Lee Sedol, mais il conserve, dit-il, la fierté d’avoir été le premier champion de Go battu par une machine. Cette compétition lui a permis de mesurer combien il est difficile de garder confiance face à une machine qui joue sans subir aucun doute, sans aucune psychologie. Cela m’a éclairé sur la capacité des machines à aider les humains à comprendre leur situation. Pour Fan Hui, AlphaGo a tout cassé, ce qui lui permet de réfléchir maintenant à la conduite du jeu, de façon beaucoup plus libre.

Assiste-t-on à un véritable progrès ? Est-ce que la nouvelle vague annoncée de l’intelligence artificielle va submerger notre intelligence ? Je ne le crois pas. Pour cela il faudrait que l’on travaille sur des algorithmes de reconnaissance du fond plutôt que de continuer à faire des algorithmes de reconnaissance des formes, mais cela suppose de se pencher sur le véritable lien qui relie syntaxe et sémantique : c’est le problème auquel on se confronte, depuis des années, tant pour le langage que pour l’intelligence artificielle.

M. Claude de Ganay, rapporteur. – Merci de ces exposés passionnants. Place, à présent, au débat interactif.

10. Débat

Mme Laurence Devillers, professeure à l'Université Paris-Sorbonne et chercheuse au Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (Limsi) du CNRS. – Il est en effet passionnant d'être confronté à tous ces défis : ils sont bien souvent les mêmes au CNRS qu'à Inria ou à l'Institut Mines-Télécom. Je vous appelle cependant, en conscience, à oublier un peu les instituts auxquels nous appartenons chacun, pour essayer de réfléchir à ce que nous pourrions faire ensemble. Si l'on veut entrer en concurrence avec les Gafa ou leur équivalent chinois, les BATX - Baidu, Alibaba, Tencent et Xiaomi -, il faut construire ensemble, autour, par exemple, d'un projet sur l'éthique et les risques numériques, on a besoin d'un institut commun de réflexion à ce sujet. En Grande-Bretagne, on a vu de nombreux projets émerger de cette façon. En France, trop de chapelles différentes coexistent, quand nous aurions besoin de bâtir une vision unifiée des efforts à mettre en œuvre. Qu'en pensez-vous ?

M. Jean Ponce, professeur et directeur du département d'informatique de l'École normale supérieure (ENS). – Beaucoup ont opposé les méthodes d'intelligence artificielle symboliques et les méthodes de type *deep learning*, attachées aux données, pour souligner le besoin d'hybridation entre l'une et l'autre. N'oublions pas, cependant, que cette hybridation existe déjà dans de nombreux domaines. Je pense, par exemple, au traitement du langage et de la parole, à la robotique, à la vision artificielle - mon domaine - où existe un effort de modélisation important : ce ne sont pas des sous-domaines de l'apprentissage statistique, qui, lui-même, ne se résume pas à l'apprentissage profond. Malheureusement, les outils classiques de l'intelligence artificielle n'ont pas eu l'impact qu'ont aujourd'hui d'autres approches.

Il a été question des besoins de financement. Le problème auquel on fait face tient au fait que nous formons d'excellents jeunes chercheurs que les grandes compagnies américaines se disputent, au point qu'attirer et retenir ces chercheurs dans nos laboratoires est extrêmement difficile, comme l'a souligné Bertrand Braunschweig, du fait des différences de salaire. Nous avons à l'École normale supérieure l'exemple des étudiants de nos deux masters spécialisés. Pour remédier à ces difficultés, il faut, à mon avis, rechercher des partenariats avec l'industrie.

M. Basile Starynkevitch, Institut List, Commissariat à l'énergie atomique (CEA). – Je suis un *geek*, j'aime coder. Je viens de terminer de rédiger une proposition qui m'a demandé trois mois de travail pour un appel qui se clôturait hier. J'ai une chance sur cent d'être retenu.

L'intelligence artificielle a encore besoin d'infrastructures logicielles. Le logiciel libre est une solution ; je suis ainsi membre de l'association April. Mais je ne suis pas un spécialiste en la matière, bien d'autres en parleraient mieux que moi.

Les projets de R&D en intelligence artificielle, parce que l'espérance de vie des start-up est de trente ou trente-cinq mois, sont le plus souvent limités à trente-six mois. Il serait souhaitable de prévoir des durées plus longues : s'il faut neuf mois à une femme pour faire un bébé, il ne faut pas un mois à neuf femmes pour en faire un. C'est la même chose pour les chercheurs !

M. Claude de Ganay, rapporteur. – Dans quelle direction les efforts de recherche en intelligence artificielle doivent-ils porter en priorité ?

M. Jean-Gabriel Ganascia. – Les financements, en France, vont à des projets de très court terme. Si les chercheurs rejoignent de grands groupes privés, c'est certes parce que les salaires y sont supérieurs, mais aussi parce que les conditions sont désastreuses dans la recherche publique : ils ont envie d'aller au bout de leurs projets, et, dans le public, cela n'est que rarement possible.

La France aurait tendance à copier les modes de financement des États-Unis ? Non, car les durées sont plus longues outre-Atlantique. Aujourd'hui, il n'y a pas de mémoire dans les institutions de financement de la recherche française...

M. David Sadek. – Un institut d'éthique de l'intelligence artificielle est une idée à explorer. Beaucoup de questions se posent, et des règles collectives seraient utiles. Nous avons besoin d'une approche inter-instituts comme le disait Laurence Devillers.

Sur l'hybridation des approches, très souhaitable également, je veux signaler qu'il y a eu, comme Jean Ponce l'a rappelé, de nombreuses tentatives, depuis longtemps, par exemple sur le traitement de la parole. On la pratique sur certains domaines. En revanche, sur les comportements plus évolués, l'interaction des machines, le dialogue, il reste beaucoup à explorer...

M. Yves Demazeau. – Une dizaine de directions de recherche ont été identifiées. Il y a trente ans, l'intelligence artificielle était une matière intégrée, autour du rêve de machine intelligente, puis on s'est rendu compte de la complexité des choses, et l'intelligence artificielle a explosé en diverses communautés spécialisées. On ne compte pas moins d'une dizaine de conférences spécialisées en intelligence artificielle. Le défi à présent est de développer une science de l'intégration. Un drone en Chine transporte un passager de cent kilos ; un véhicule autonome se déplace dans les rues de villes américaines. Mais ni l'un ni l'autre ne rassemble les meilleures technologies des sous-disciplines. Il reste à les réunir !

Mme Dominique Gillot, rapporteure. – Je perçois dans vos propos, pourtant, le souci d'une démarche partagée, l'aspiration à l'interdisciplinarité. C'est un enjeu dont vous avez conscience.

M. Yves Demazeau. – Entre chercheurs, oui, mais pas entre organismes de recherche, ni entre institutions ou acteurs du secteur.

M. Jean-Daniel Kant. – L'offre, dans l'enseignement, est peut-être insuffisante, mais je signale que mon université propose deux masters en intelligence artificielle, l'un orienté sur les données, l'autre sur les comportements. Ils comptent au plus quatre-vingts étudiants, mais ils existent !

L'hybridation doit-elle être une priorité ? Pas nécessairement, car elle ne sera pas facile à faire, en tout cas sur les fonctions de haut niveau où ce sera très difficile. C'est facile sur des fonctions de bas niveau. On a essayé il y a une vingtaine d'années de plaquer des symboles dans les neurones, sans résultat. On peut aussi faire l'inverse... Il faut réfléchir avec les disciplines concernées et reprendre la question, sur l'apprentissage par exemple.

Mme Laurence Devillers. – L'hybridation existe déjà. Je vois le cas des dialogues hommes-machines. Si l'on veut être compétitif, il faut éviter de se poser éternellement des questions théoriques primaires.

M. Igor Carron. – Je suis le P-DG d'une petite start-up de *hardware* pour le *machine learning*. Je fais également partie d'un groupe qui organise les *Paris Learning Machine Meetups*, avec à ce jour soixante *meetups*, durant lesquels nous avons accueilli deux cents intervenants, autour d'un réseau de plus de 4 700 personnes. Un autre monde existe en matière d'intelligence artificielle : des personnes sorties d'école depuis un, deux ou trente ans se réunissent, non dans le cadre universitaire, mais au contact direct des recherches en cours, pour comprendre ce qui se passe.

Le besoin de l'État n'est pas de même nature en France qu'aux États-Unis : les discussions de ce matin sont très intéressantes, car l'aspect *push* est important ; on a envie de partager, d'échanger, d'envisager des projets. Mais l'aspect *pull* fait défaut en France. En Amérique, chaque administration a son programme pour associer start-up et grands groupes à la recherche du futur, sur la sécurité, sur les risques, etc. C'est le problème des incitations.

Bien des *a priori* qui existaient il y a vingt ans ont disparu, la connexion entre réseaux de neurones et approches symboliques progresse, l'explication des modèles aussi. Je suis allé au *Nips* – la *Conference on Neural Information Processing Systems*, le Davos de l'intelligence artificielle – à Barcelone récemment : j'ai constaté la faiblesse de la représentation française dans certains *workshops* absolument primordiaux, comme l'interprétation des modèles ou les composantes éthiques, par exemple.

Pourquoi n'y a-t-il pas en France une politique de *pull* assumée par l'administration, qui interrogerait les chercheurs sur certains sujets d'avenir ?

Mme Dominique Gillot, rapporteure. – Il y a deux ans que je sollicite l'OPECST pour engager ce travail. Hélas ! C'est dans l'urgence, juste avant la fin de nos mandats parlementaires, que nous nous penchons sur le sujet. Le secrétaire d'État à l'enseignement supérieur et la secrétaire d'État au numérique présenteront demain le début d'une stratégie nationale pour la recherche en intelligence artificielle : répondez à ces invitations, inscrivez-vous dans les ateliers proposés ! Le Conseil national de stratégie de la recherche vient d'étudier le livre blanc qui clôture la séquence des cinq dernières années, en matière d'enseignement supérieur et de recherche – la recherche y apparaît bien comme une nécessité fondamentale, la liberté des chercheurs étant mise en avant.

J'ai l'impression que les politiques sont toujours en réponse, jamais en tête de pont – ou en *push* ou, mieux, en *pull*, pour reprendre vos termes. C'est vrai que le politique n'est pas en *pull*. C'est bien là l'enjeu de rencontres comme celle d'aujourd'hui, qui doivent être valorisées, afin que les chercheurs, ceux qui travaillent sur l'avenir, orientent la décision politique. Aux États-Unis, les décideurs ont manifestement plus d'audace, ils prennent des risques, alors que, chez nous, en politique comme ailleurs, le principe de précaution l'emporte...

M. Jean Ponce. – Je veux dire, après avoir écouté M. Demazeau, qu'il existe un bon domaine intégrateur pour l'intelligence artificielle, c'est la robotique. Il doit être poussé et valorisé comme tel. Il y a quelques jours, aux États-Unis, deux cent cinquante millions de dollars ont été débloqués autour de la Carnegie Mellon University, à Pittsburgh, par un effort public-privé, pour financer des projets de robotique.

M. Patrick Aknin. – Je suis directeur scientifique de l'Institut de recherche technologique SystemX, qui fait le lien entre laboratoires et industrie sur des sujets où intervient l'intelligence artificielle, comme les véhicules autonomes, les *smart grids* ou réseaux électriques intelligents, les territoires intelligents, etc. Un frein à l'exploitation de l'intelligence artificielle réside, me semble-t-il, dans les lacunes de l'explication des décisions automatisées, leur « prouvabilité ». Il y a là un défi sociétal. S'y pencher permettrait de déverrouiller des propositions de chercheurs, afin d'aider l'homme dans sa vie quotidienne.

Cédric Villani, qui n'est pas connu comme un pro-connectionniste ou un mentor de l'apprentissage statistique, a parlé pourtant d'interactions dans le secteur : certains chercheurs assurent que, dans quelques années, on aura des preuves et une théorisation de ces approches par apprentissage, par exemple avec la théorie de *Vapnik* qui tente d'expliquer l'apprentissage d'un point de vue statistique. Après l'émergence du *deep learning* dans la période

actuelle, on aura des propositions de théorisation. Il faudrait encourager ce type de travaux, pour dépasser les peurs collectives sur ces sujets.

M. Patrick Albert. – Je suis l'un des vétérans dans ce domaine, la *good old-fashioned AI* ou bonne IA à l'ancienne, avec une thèse de doctorat au centre de recherche de Bull, la création d'un premier centre industriel puis une des premières start-up, ILOG, dont le nom vient de l'abréviation des termes intelligence logicielle... vendue, hélas !, à IBM, car il n'y avait personne pour la racheter en France. Bref, j'ai une carrière à la fois scientifique et industrielle, et je suis membre du bureau de l'Association française pour l'intelligence artificielle (AFIA).

Pour déterminer des priorités, les critères sont nombreux, scientifiques, économiques, ayant trait à la défense nationale, à la citoyenneté, etc. Nous vivons de plus en plus sous le règne de l'algocratie, avec des algorithmes qui intègrent un apprentissage et qui prennent des décisions, mais sur quels motifs ? Google et Facebook sont ainsi régulièrement accusés de racisme : ils se bornent à répondre que leurs systèmes créent des distributions de probabilité qui fonctionnent. Ce n'est pas socialement acceptable, il faut pouvoir expliquer les décisions. Les entreprises doivent pouvoir justifier leurs choix. Pour l'allocation des prêts, les banques utilisent beaucoup ces outils, mais elles savent expliquer en partie la décision prise.

L'organisation de la recherche est un sujet difficile, il faudra pourtant progresser sur ce point pour mieux développer l'intelligence artificielle. J'en parle d'autant plus librement que je n'appartiens à aucune organisation. Chaque sous-domaine de ce secteur de la recherche a des velléités d'indépendance ; il est compliqué de réunir toutes les communautés. Des efforts sont à accomplir. L'intelligence artificielle est une technologie diffusante, dont les impacts doivent être pris en considération. Il faut aider les différentes communautés à s'organiser. Une organisation la moins centralisée possible serait la meilleure forme, même s'il faut fédérer l'ensemble des travaux et représenter le secteur auprès de l'État et la société.

M. Claude de Ganay, rapporteur. – Je vous remercie tous de votre participation à cette première table ronde et je laisse Dominique Gillot présider la deuxième table ronde.

III. DEUXIÈME TABLE RONDE, PRÉSIDIÉE PAR MME DOMINIQUE GILLOT, RAPPORTEURE : DIMENSIONS STRATÉGIQUES EN MATIÈRE DE RECHERCHE EN INTELLIGENCE ARTIFICIELLE

Mme Dominique Gillot, rapporteure. - J'invite en premier lieu Jean-Marc Merriaux, directeur général de Canopé, qui est la prolongation du Centre national de documentation pédagogique (CNDP). Après avoir effectué une carrière de quinze ans à France 5, Jean-Marc Merriaux est toujours très intéressé par les actions destinées à la communauté éducative et s'investit dans l'éducation populaire. Vous avez sept minutes pour dresser un portrait des enjeux éducatifs de l'intelligence artificielle.

1. M. Jean-Marc Merriaux, directeur général de Canopé

Merci pour cette invitation. Bien évidemment, je ne suis ni un spécialiste de l'intelligence artificielle ni un chercheur. Je suis plutôt ici à titre de médiateur faisant le lien entre tous les enjeux pédagogiques attachés aux évolutions technologiques. Canopé, opérateur du ministère de l'éducation nationale, est présent pour accompagner l'évolution des pratiques pédagogiques.

Enjeux de l'intelligence artificielle

Interrogé sur la prise en charge de l'intelligence artificielle par l'éducation, je me suis efforcé de bien positionner les thèmes du débat, à commencer par la problématique majeure de l'incommunicabilité entre les sciences, les technologies et la société.

L'éducation vise à repositionner l'humain au cœur de l'ensemble des dispositifs. L'humain peut-il changer grâce et par le numérique ? L'autre problématique a trait à la place de l'humain dans une société hyperconnectée : robots, transhumanisme, interactions avec les machines, ubérisation des économies... Nous nous situons également dans un rapport au temps entre la pédagogie et l'ensemble de ces questions technologiques. Malgré le fait que les sciences et technologies n'ont jamais autant accompli pour améliorer la vie de l'homme, l'ensemble de ces technologies nouvelles font peur, ce qu'il ne faut pas sous-estimer.

L'intelligence artificielle aujourd'hui, sait comprendre ce qu'on lui demande, s'exprimer correctement et répondre à des questions simples. Nous avons évoqué Siri et un certain nombre d'autres évolutions lors de la première table ronde. Néanmoins l'intelligence artificielle ne sait pas encore comprendre le sens du langage adapté au contexte. On sait que le sens du

langage et de l'oralité sont intégrés de manière très forte dans le cadre pédagogique. Demain, l'intelligence artificielle pourra sans doute répondre à des enjeux pédagogiques et inventer son propre langage, expérimenter le monde pour comprendre le sens du langage. L'éducation est aujourd'hui notre seul atout pour réduire l'incommunication entre sciences, technologie et société. Les pédagogues, au premier chef, doivent en avoir pleinement conscience. De ce fait, se pose la question de la place des sciences et des technologies du numérique dans tous les apprentissages pour comprendre une société qui se complexifie.

Apports pédagogiques de l'intelligence artificielle

Sur les apports pédagogiques de l'intelligence artificielle, j'ai tenté de dresser un éclairage rapide. De nouvelles pratiques pédagogiques pourraient émerger en dépassant la relation émetteur-récepteur et en transformant l'évaluation. Il est vrai que les enjeux de l'intelligence artificielle nous conduisent à nous interroger sur l'évaluation et, entre autres, sur les moyens de prédire la réussite des élèves. Il nous faut pouvoir intégrer cette dimension. De plus, la répétition est avant tout un acte pédagogique, de sorte que l'intelligence artificielle repose beaucoup sur ces activités de répétition. Par ailleurs, la différenciation des apprentissages constitue un autre aspect essentiel de la pédagogie, la personnalisation devant être adaptée à la diversité des élèves. J'insisterai enfin sur un dernier point, tenant au *continuum* pédagogique entre le temps scolaire et le hors temps scolaire. La présence future, peut-être, de robots aussi bien à l'école qu'au sein de la maison, devra être accompagnée d'outils et d'interfaces pour assurer les usages au sein et en dehors de la classe. Sur ce point, l'intelligence artificielle peut sûrement apporter un certain nombre de réponses.

J'insisterai essentiellement sur la place et le rôle de l'enseignant du fait de l'ensemble des évolutions évoquées, ainsi que sur la nécessité de l'accompagner dès lors que les nouvelles technologies seront parties intégrantes de la classe. L'intelligence artificielle va transformer le geste professionnel de l'enseignant. De même, la médiation constituera un outil important pour répondre aux enjeux technologiques. Selon moi, les nouvelles technologies ne sont pas conçues pour être en compétition avec les enseignants. L'intelligence artificielle intervient surtout pour compléter le savoir-faire de l'enseignant, en le rendant plus accessible et mieux informé. La posture du médiateur constitue par conséquent un enjeu essentiel de la nouvelle relation à l'élève. La médiation représente un enjeu de réassurance pour l'éducation dans un monde en pleine mutation technologique. Elle renforce la place de l'humain et garantit également le collectif. Le médiateur est un tiers de confiance, ce qualificatif s'appliquant *a fortiori* à l'enseignant. La médiation favorise en outre une meilleure adéquation avec les besoins de chaque usager.

Enfin, j'évoquerai un exemple sur lequel je travaille dans le cadre du projet *e-fran*, qui liera aussi bien des pédagogues que des chercheurs et des

start-uppers. Nous avons développé Matador, qui était à la base un jeu de plateau mais qui a été transformé dans un environnement numérique. Le projet repose sur un monitoring individuel d'apprentissage du calcul mental par chaque élève. Plus l'élève jouera en classe et à la maison et mieux l'enseignant connaîtra ses compétences acquises et non acquises. Nous travaillerons sur le parcours d'un élève qui sera mis en rapport avec tous les autres élèves de son niveau scolaire. L'interaction et l'horizontalité constituent par conséquent des éléments importants de ce type de projet. *E-fran* reposera sur mille cinq cents élèves pendant une année scolaire, avec l'objectif d'analyser plus de sept cent mille opérations chaque année. Selon les profils, des parcours de jeu spécifiques seront proposés à chaque élève. L'ensemble s'appuie sur des chercheurs, aussi bien statisticiens que didacticiens et cognitivistes. Nous travaillons dans cette dynamique en intégrant l'ensemble des questions développées à l'occasion de la première table ronde, même si nos enjeux ne sont pas aussi poussés que certains domaines de l'intelligence artificielle.

Je suis au regret de devoir m'éclipser pour prendre un train. Je peux répondre aux questions maintenant.

Mme Dominique Gillot, rapporteure. – Nous devons préalablement achever les exposés. Je passe donc la parole à François Taddéi, qui possède de multiples compétences mais est invité ce jour en sa qualité de directeur du CRI, donc praticien des rencontres des intelligences. Son intervention portera sur le travail qu'il a engagé concernant l'évolution des pratiques pédagogiques liée à l'évolution des technologies.

2. M. François Taddei, directeur du Centre de recherche interdisciplinaire (CRI)

Effectivement, la ministre de l'Éducation nationale m'a confié une mission ayant trait à la société apprenante. Il s'agit de déterminer comment nous apprenons tout au long de la vie, que ce soit dans la petite enfance, dans l'enseignement scolaire supérieur ou au cours de la vie professionnelle.

Aujourd'hui, l'intelligence artificielle et l'intelligence humaine co-évoluent. En ma qualité de biologiste de l'évolution, je constate que nous avons connu un certain nombre de grandes transitions évolutives avec l'apparition de la vie et même avec celle des cerveaux. La connexion de ces cerveaux entre eux délivre une série de messages coordonnés pour progressivement accomplir un nombre croissant de réalisations. Puis l'homme est apparu et a inventé d'autres manières, toujours plus élaborées, de communiquer. Le langage est bien entendu l'une d'entre elles, avant l'écriture, l'imprimerie et l'informatique. Il existe aujourd'hui une accélération de ces transitions qui comportent une certaine part de singularité technologique. Dès lors, nous avons besoin d'inviter des acteurs

de tous horizons à réfléchir à cette pluralité de futurs, que l'on peut souhaiter ou non. Cette pluralité d'acteurs inclut les enfants, toujours très intéressés par les thématiques nouvelles. Lorsqu'on donne aux enfants la possibilité de s'interroger, nous constatons qu'ils peuvent aller très loin. Nous avons ainsi développé un programme dénommé « Les savanturiers », à l'occasion duquel des enfants se questionnent et travaillent avec des chercheurs. Assez spontanément, ils posent des questions telles que l'existence avant le *big bang* ou le devenir de l'humain. Les enfants sont fondamentalement attirés par les sujets ayant trait à l'intelligence artificielle et la robotique. Ils se demandent spontanément quelle sera leur place dans ce monde et s'ils co-évolueront avec les nouvelles formes d'intelligence.

L'intelligence collective est présente depuis longtemps dans le vivant, mais passe progressivement à un collectif d'hommes et de machines. Dans le monde des échecs, on parle d'IA-centaures, combinaisons d'hommes et de machines qui battent les hommes mais aussi les machines jouant seuls. Voilà ce qui marche le mieux.

Le besoin de questionnement éthique est en outre présent. Il nous interroge sur la place de l'homme. Les plateformes de recherche participative peuvent mobiliser l'intelligence collective au service de grands enjeux. François Houllier a ainsi rendu plusieurs rapports au terme de la mission sur les sciences participatives qui lui a été confiée. Les intelligences artificielles sont capables d'optimiser un jeu de données mais ne peuvent pas spontanément inventer un autre jeu ou se donner d'autres objectifs. L'école cherche à apprendre à mémoriser le passé, ce que l'intelligence artificielle sait faire, mais il faut aussi inventer. Cette capacité à se questionner est encore une spécificité humaine qu'il faut approfondir toujours plus. Nous devons travailler sur la complémentarité, coopérer pour questionner demain et co-construire ces solutions. Les grands défis planétaires nécessitent toujours plus d'intelligence. Il faut convaincre nos concitoyens que nous ne devons pas arrêter la course, mais plutôt réfléchir aux raisons pour lesquelles nous courons. Nous devons nous poser ces questions collectivement, en ne laissant pas l'évolution technologique s'affranchir des questions éthiques. Aristote disait déjà qu'il existe trois formes de connaissance : *épistémè*, la connaissance du monde (qui a donné l'épistémologie et la science), *technè*, l'action sur le monde (qui a entraîné l'apparition des technologies dont nous discutons aujourd'hui) et *phronésis*, l'éthique de l'action ou sagesse pratique. Nous avons donc besoin de réfléchir simultanément à la coévolution de ces sciences technologiques et de cette éthique individuelle et collective. La réflexion doit être de court et long termes, locale et globale. C'est pourquoi il est nécessaire de former les décideurs sur les grands enjeux. Par conséquent, je me félicite de cette mission parlementaire. La prise de conscience de ces enjeux doit faire l'objet d'un débat toujours plus large et évidemment international. Nous devons impliquer toujours plus d'acteurs dans la réflexion. Notre système éducatif et notre enseignement supérieur devront être pensés pour s'adapter au monde de demain. Les institutions

d'intelligence humaine seront tenues de se réinventer à l'heure où les machines vont les concurrencer.

Mme Dominique Gillot, rapporteure. – La place et la transmission de la connaissance constituent évidemment des enjeux fondamentaux. Je vais passer la parole à Olivier Esper, responsable des affaires publiques de Google France. Vous allez nous présenter les activités de Google et de sa filiale en matière de recherche en intelligence artificielle. Nous sommes évidemment tous très intéressés par ce qui se passe chez Google, notamment par l'appel à la sensibilisation à la responsabilité sociétale de cette grande société.

3. M. Olivier Esper, responsable des affaires publiques de Google France

Je vous remercie de m'avoir invité à vos travaux pour y apporter ma modeste contribution, n'étant moi-même pas un chercheur. Je sais que vous prévoyez également de vous déplacer à l'international, j'espère que vous aurez l'occasion de rencontrer Greg Corrado le directeur de la recherche de Google.

Mme Dominique Gillot, rapporteure. – Oui c'est bien prévu lors de notre déplacement aux États-Unis la semaine prochaine.

M. Olivier Esper. – Très bien. À l'été 2015, les fondateurs de Google ont créé une maison-mère nommée Alphabet pour permettre à différentes entreprises distinctes - et désormais autonomes par rapport à Google - de porter des projets intégrant de l'intelligence artificielle : *Waymo* pour la voiture autonome, *Deep Mind* pour le programme AlphaGo...

Pour sa part, Google concentre essentiellement ses recherches sur l'apprentissage automatique, le *machine learning*, pour concevoir des machines qui répondront intelligemment à des solutions et situations nouvelles. Ces machines sont entraînées avec des jeux de données. La recherche menée par Google correspond à un modèle ouvert, que des ingénieurs extérieurs à Google peuvent utiliser dans d'autres projets. Le constat posé par Google est celui du franchissement actuel d'un cap. De plus en plus d'usages, de traductions concrètes de la recherche existent aujourd'hui dans le monde réel, et sont la résultante de capacités de calcul augmentées, de sources d'informations très riches et de communautés larges de chercheurs et d'ingénieurs.

La plupart des services de Google intègrent de l'apprentissage machine, avec trois objectifs. Le premier d'entre eux est d'améliorer la performance des produits. Par exemple, *Google Translate*, appelé en France *Google Traduction*, reposait initialement sur une traduction mot à mot. En 2016, la nouvelle version de ce service intègre de l'apprentissage machine. Ce changement a abouti à une augmentation des performances inégalée au cours des 10 dernières années. Le second de ces objectifs est de proposer des

nouvelles fonctionnalités. Par exemple *GooglePhoto* permet désormais aux utilisateurs d'effectuer des recherches dans leur bibliothèque sur des mots-clés sans avoir au préalable annoté ces photos. Le troisième objectif des chercheurs de Google est de se confronter à des problématiques à grande échelle. Par exemple, Gmail est maintenant capable de bloquer plus de 99,9 % du spam, y compris des variantes de messages jamais encore vus auparavant, grâce à l'utilisation de techniques de réseaux neuronaux.

En dépit des progrès réalisés, les champs de recherche sont encore nombreux. J'en citerai trois identifiés par nos ingénieurs, et qui feront écho à certains éléments que nous avons abordés précédemment. Le premier champ de recherche a trait à la compréhension des modèles. Il s'agit de développer des modèles auto-explicatifs susceptibles d'expliquer un résultat par rapport à une situation spécifique. En deuxième lieu, les ingénieurs Google travaillent sur l'apprentissage machine à partir de petits jeux de données. En la matière la qualité et le volume de ces données sont d'importance. Par exemple, pour reconnaître un chat sur une photo, nous avons besoin de millions de photos de chats pour entraîner les machines. Or la possibilité d'entraîner ces machines sur de plus petits jeux d'échantillons permettrait de mettre en œuvre l'apprentissage automatique sur un plus grand nombre de domaines.

Enfin, le troisième champ de recherche concerne l'amélioration de la compréhension de sémantiques subtiles des modèles. Pour reprendre l'exemple des photos, le modèle est capable de repérer un chat mais ne sait pas encore dire si la photo est triste ou drôle. Une telle contextualisation de l'image n'est pas encore permise. De même, le système ne sait pas encore comprendre qu'une vache volant dans le ciel n'est pas une situation normale.

En parallèle de ces domaines de recherche de nature scientifique, des travaux sont également réalisés sur des enjeux sociétaux et éthiques. Comme d'autres technologies, l'apprentissage machine n'est pas bon ou mauvais en soi car tout dépend de l'usage qui en est fait. Afin de réfléchir aux enjeux sociétaux éthiques, nous sommes membres fondateurs de « *Partnership on AI* »¹ avec Facebook et d'autres entreprises, afin de développer des bonnes pratiques et de maximiser les retombées positives des nouvelles technologies. Ici encore, je citerai trois domaines sur lesquels nous travaillons. En premier lieu sur les risques de sécurité que pourrait poser l'intelligence artificielle, nos chercheurs ont publié l'an dernier un papier en collaboration avec les universités de Stanford et de Berkeley².

Le deuxième exemple de recherches a trait au risque de partialité ou de discrimination si la machine est elle-même entraînée sur un jeu de données biaisé. Dans ce domaine encore, nos ingénieurs ont publié des travaux. Lors de la conférence NIPS à Barcelone, organisée en décembre

¹ <https://www.partnershiponai.org>

² <https://research.googleblog.com/2016/06/bringing-precision-to-ai-safety.html>

dernier sur l'intelligence artificielle, deux de nos ingénieurs ont publié un papier pour cerner la problématique et proposer une approche de solutions¹.

En troisième lieu, l'impact économique de l'intelligence artificielle, c'est-à-dire sa traduction concrète sur les emplois présents et futurs, conduit à réfléchir aux métiers de demain et à ajuster l'éducation en conséquence. Dans ce domaine, les études sont loin de converger vers les mêmes chiffres, mais il importe d'ores et déjà de réfléchir à la formation.

En conclusion au-delà des exemples que j'ai mentionnés, l'objectif de Google est de rendre ses services plus performants et utiles dans la vie quotidienne des gens. L'intelligence artificielle permettra aussi de se confronter aux grands problèmes de notre société. Dans le domaine de la lutte contre le réchauffement climatique, Google a utilisé de l'apprentissage machine pour optimiser l'allocation de ressources de ses *data centers*. Cette mesure a permis une réduction de 40 % de la consommation d'énergie. L'entreprise s'est en effet fixé pour objectif de parvenir dès la fin 2017 à 100 % d'approvisionnement en énergie renouvelable, en particulier grâce à l'utilisation de l'intelligence artificielle. Ces avancées technologiques ont donc un potentiel positif.

Mme Dominique Gillot, rapporteure. - Delphine Reyre est directrice Europe des affaires publiques de Facebook, deuxième site *web* le plus visité au monde après Google. Facebook fait un usage croissant des technologies d'intelligence artificielle. J'ai rencontré à deux reprises Yann LeCun, directeur de la recherche en intelligence artificielle de Facebook, qui dirige la recherche mondiale de Facebook mais a installé un laboratoire de recherche en France.

4. Mme Delphine Reyre, directrice Europe des affaires publiques de Facebook

Merci beaucoup de votre invitation. Je souhaite juste apporter une petite note personnelle. Avant de rejoindre Facebook, j'ai travaillé pour un grand ministre de la recherche, Hubert Curien, qui entretenait des liens très étroits avec l'OPECST. Je suis donc présente parmi vous avec un plaisir particulier. Olivier a très bien décrit comment les grandes entreprises du numérique abordent les sujets de l'intelligence artificielle. Je ne serai par conséquent pas trop redondante, étant précisé que Yann LeCun sera disposé à répondre à toutes les autres questions par la suite. Je vous propose plutôt de vous apporter un éclairage particulier, car sous l'impulsion de Yann nous avons choisi d'établir notre laboratoire d'intelligence artificielle en grande partie à Paris. Je préciserai comment s'exerce le dialogue entre la recherche publique et la recherche privée sur ces domaines en évolution rapide.

¹ <https://research.googleblog.com/2016/10/equality-of-opportunity-in-machine.html>

Facebook a investi deux grands domaines d'intelligence artificielle. Un groupe travaille sur la recherche appliquée, en particulier sur des applications concrètes dont l'un des exemples-types concerne les capacités de reconnaissance visuelle. Ces applications permettent à des malvoyants ou des non-voyants d'utiliser Facebook et donc d'accéder au monde des réseaux sociaux. La machine doit par exemple permettre de lire ou de décrire à ces personnes l'image qu'ils reçoivent sur leur fil d'actualité. Le décloisonnement des usages des médias sociaux est ainsi opéré au bénéfice des personnes atteintes de handicaps. Dans son ensemble, le thème de l'accessibilité connaîtra un développement rapide.

Par ailleurs, je souhaitais évoquer plus précisément notre activité de recherche fondamentale, dénommée Facebook *Artificial Intelligence Research (FAIR)*, développée par trois laboratoires basés respectivement à Menlo Park dans la Silicon Valley, New York où enseigne Yann LeCun, et Paris. Le choix de Paris s'est bien entendu opéré sous l'impulsion de ce grand chercheur qui a bien voulu travailler avec Facebook, mais aussi parce que l'excellence de l'école mathématique française est mondialement reconnue. Inria, l'ENS, l'Université Pierre-et-Marie-Curie ainsi que d'autres universités sont véritablement des sources d'excellence, de sorte qu'il nous a paru important de pouvoir travailler avec leurs chercheurs sans nécessairement les faire partir aux États-Unis. D'ailleurs ces chercheurs ne souhaitaient pas véritablement s'expatrier. Ainsi, l'installation du laboratoire à Paris nous a permis de fixer les chercheurs en France, d'en faire revenir d'autres et même d'attirer des chercheurs étrangers. Sur vingt-cinq chercheurs accueillis à Paris, figure en effet une proportion importante de chercheurs étrangers. Telle est la preuve que la France peut être un centre d'excellence et une terre d'attraction pour l'intelligence artificielle.

Ainsi que je l'évoquais, nous travaillons avec la recherche publique et en partenariat étroit avec Inria. Je souhaite également apporter un éclairage sur un instrument très spécifique à la France, les Conventions industrielles de formation par la recherche (CIFRE), qui permettent une interaction formidable entre la recherche publique et la recherche privée. Les thésards et les doctorants accueillis en milieu de recherche privée, auront sans doute vocation à exercer dans la recherche publique. Facebook a également mis en place un programme de dons de moyens de calcul à des universités françaises et européennes, parmi lesquelles Inria, l'ENS et l'Université Pierre-et-Marie-Curie. Nous avons récemment annoncé un certain nombre de dons de ce type en France.

Lorsque nous maintenons des chercheurs en France, ils continuent à participer activement à la communauté scientifique française. Ces chercheurs enseignent et apportent leur contribution à des conférences. Yann LeCun a, à ce titre, ressenti comme un très grand honneur de donner des conférences au Collège de France, dans le cadre de sa chaire d'informatique.

Enfin, la quasi-totalité des travaux effectués par Facebook sont publiés en mode ouvert, de sorte qu'ils sont immédiatement saisissables par

tous chercheurs. Il existe donc un véritable effet de démocratisation en intelligence artificielle grâce à ce mode ouvert. La recherche n'est pas réservée à de grands groupes privés et est mise à la disposition de tous.

En termes d'éthique, Facebook est partenaire avec Microsoft, Google, IBM et Amazon du *Partnership on AI*. Il est également important de souligner que nous travaillons quasi exclusivement sur des données publiques.

Mme Dominique Gillot, rapporteure. - Vous n'avez pas parlé du crédit d'impôt recherche (CIR).

Mme Delphine Reyre. - Notre objectif est de trouver avant tout des cerveaux. À cet égard le CIFRE nous satisfait davantage.

Mme Dominique Gillot, rapporteure. - Merci beaucoup pour votre exposé très précis. L'intervenant suivant est Laurent Massoulié. Votre expérience est également très intéressante puisque vous êtes directeur d'un centre de recherche commun à Inria et à Microsoft. Ce type de partenariat a vraisemblablement vocation à se développer.

5. M. Laurent Massoulié, directeur du Centre de recherche commun Inria-Microsoft

Je vous remercie de votre invitation. J'ai rejoint Inria il y a quatre ans pour diriger le laboratoire commun entre Microsoft Research et Inria. Je suis chercheur dans les réseaux et les systèmes distribués. Le laboratoire commun Microsoft Research-Inria est né il y a dix ans d'une ambition de collaboration étroite entre les directeurs de recherche de Microsoft et d'Inria. Le projet a été rendu possible car les chercheurs se connaissaient. Actuellement, le laboratoire héberge une quarantaine de chercheurs travaillant sur six projets, dont quatre sont liés à l'intelligence artificielle.

L'un de ces projets, ayant trait à la vision sur ordinateur et dirigé par Jean Ponce ici présent, traite de la compréhension automatique des vidéos. Un autre projet concerne l'imagerie médicale, avec l'objectif de parvenir à poser des diagnostics personnalisés et à aboutir à une meilleure prévention grâce à l'intelligence artificielle. Enfin nous avons des projets autour des fondamentaux des algorithmes pour l'apprentissage supervisé ou non supervisé, notamment dans le *cloud*, par exemple, la plateforme de *cloud computing* de Microsoft.

Force est de constater que les partenariats entre un institut tel qu'Inria et les laboratoires de recherche industrielle comme celui de Microsoft sont non seulement faisables, mais seront de plus en plus nombreux à l'avenir. En réalité, les chercheurs qui font avancer la science dans le privé et le public sont issus des mêmes formations, se connaissent et assistent aux mêmes conférences. Par conséquent, la frontière entre le privé et le public est de plus en plus perméable, et ce dans les deux sens. Dans la

conduite de nos projets, nous avons formé de jeunes chercheurs qui, pour un certain nombre d'entre eux, ont été recrutés ensuite par Microsoft. Nous avons également connu des transferts dans l'autre direction. Pour ma part, je suis l'un des trois cas de chercheurs ayant quitté Microsoft pour rejoindre Inria.

L'objectif des partenariats public-privé vise en premier lieu à ce que le privé apporte ses problèmes à la recherche publique. Les chercheurs en vision tels que ceux de l'équipe de Jean Ponce, s'ils peuvent être au contact avec les équipes de Microsoft sur la réalité augmentée, seront ses contributeurs sur le sujet. Dans l'autre sens, il existe un engouement des grands groupes tels que Microsoft sur le fait de travailler avec les chercheurs d'Inria, dont l'excellence est reconnue et qui disposent d'une grande latitude pour s'engager sur des objectifs de long terme. Par exemple lorsque Microsoft décide de s'intéresser à l'imagerie médicale, les chercheurs d'Inria recrutés ont déjà des dialogues approfondis avec les centres hospitaliers universitaires (CHU) et sont en prise avec les vrais problèmes de l'imagerie médicale.

Je conclurai en affirmant que le partenariat public-privé est un excellent moyen de faire avancer la connaissance dans le privé et l'innovation dans le public. De plus l'attractivité des salaires dans le public restant moindre que dans le privé, ce genre de partenariats peut représenter une solution.

Mme Dominique Gillot, rapporteure. – Sans transition, je donne la parole à Dominique Cardon, professeur de sociologie à l'Institut d'Études Politiques de Paris et ancien de Médialab, spécialisé dans la sociologie de la recherche et des mondes numériques.

6. M. Dominique Cardon, professeur de sociologie à l'Institut d'Études Politiques de Paris/Médialab

L'intelligence artificielle est née trois fois, morte deux fois, et réapparaît aujourd'hui dans l'espace public. Elle a sans doute un problème de communication. Pour l'apprivoiser, nous aurions besoin de nous défaire de cette hypothèse d'estompement de la frontière entre l'humain et l'inhumain, et de cette obsession anthropomorphique. L'intelligence artificielle revêt plusieurs formes car les conceptions de l'intelligence sont différentes, il y a donc plusieurs types d'intelligence artificielle. J'observe une tentative des anciennes générations de prendre le train des nouvelles formes d'intelligence artificielle, comme le *deep learning*. Il est important pour la collectivité de chercheurs de ne pas faire de promesses inconsidérées dans cette représentation de la science autonome de la machine.

Par ailleurs, les nouvelles techniques étant massivement statistiques, la question des données est très centrale. Les enjeux posés à la société sont à la fois éthiques et concernent la complémentarité entre recherche publique et

recherche privée. Il faut se demander à partir de quelles données les modèles sont aujourd'hui construits. Les données appartiennent en effet aux grandes entreprises privées, les GAFAs en particulier (Google, Apple, Facebook, Amazon). On voit apparaître de nouvelles formes d'articulation entre recherche publique et recherche privée, sous l'impulsion des *meetups* en particulier, comme celui qu'organise Igor Carron

Comprendre, critiquer et apprivoiser sont les principaux enjeux. Le premier de ces enjeux, épistémologique, vise à comprendre le fonctionnement de l'intelligence artificielle et à éduquer la société aux nouveaux objets. Il faut en faire à la fois un enjeu scientifique et éducatif, ainsi qu'il l'a été rappelé. À ce titre, il est nécessaire d'étudier l'informatique, les mathématiques et les statistiques. Google possède une manière de classer l'information qui lui est propre, Facebook a également une pratique spécifique, elle aussi discutable. Il est par conséquent nécessaire de faire comprendre à nos sociétés les visions du monde qui sont derrière ces représentations pour savoir les critiquer. Les inquiétudes sur ces questions sont nombreuses. Il existe des biais.

La dynamique actuelle autour du Conseil national du Numérique et d'Inria correspond à un besoin de notre société de se saisir des enjeux des algorithmes, en dialogue avec les chercheurs, pour éventuellement vérifier les classements d'un certain nombre d'opérations ou le respect de principes de diversité. Il faut qu'il y ait des débats. Sur les monopoles, sur la fixation des prix, sur le respect du principe de diversité ou sur les bulles d'informations, il faut que la société soit partie prenante de ces discussions. Le concept de loyauté des algorithmes vise à mettre en place des techniques informatiques pour vérifier la manière dont les utilisateurs comprennent ce que fait l'algorithme pour eux. Les travaux de Benjamin Edelman à ce sujet méritent d'être cités.

Apprivoiser les différentes formes d'intelligence artificielle suppose de cesser de penser les intelligences artificielles comme des machines autonomes. C'est une image dont il faut se défaire car nous en sommes encore très loin en vérité. Il faut se désenvoûter de ces croyances pour mieux vivre avec ces nouvelles technologies.

Mme Dominique Gillot, rapporteure. – Merci d'avoir posé la question de la détention des données, qui suscite beaucoup de fantasmes, et de la capacité à les engranger. Le monopole pose question à notre société et à nos usagers. Nous terminons avec Gilles Babinet, dont le témoignage est précieux car il a été dès 2011, le premier président du Conseil National du Numérique et il est aujourd'hui *Digital Champion* auprès de la Commission européenne, il pourra profiter de son intervention pour nous parler de cette nouvelle mission.

7. M. Gilles Babinet, entrepreneur, *digital champion* auprès de la Commission européenne

Bonjour, et merci de votre invitation à participer à vos travaux. J'ai quelque peu restructuré mes propos au fur et à mesure des interventions de ce matin. Je ne possède aucune expertise poussée en intelligence artificielle, même si ce sujet me passionne en tant que néophyte. Pour autant, je souhaite le rapprocher des enjeux productifs et d'enseignement en général, aussi bien dans le primaire que dans le supérieur. Depuis une dizaine d'années, une résurgence de l'intelligence artificielle a été constatée, notamment sous l'impulsion de personnalités telles que Yann LeCun et d'autres encore. Les applications concrètes sont disponibles depuis peu. Des intelligences artificielles sont déjà partout, dans notre quotidien. Nous avons tous été confrontés à une intelligence artificielle *via* nos *smartphones* ou les feux rouges dans les villes. Cette intelligence artificielle faible, qui fait partie du quotidien de chacun de nous, devrait à ce titre être davantage enseignée et être davantage connue.

L'impact productif constitue une préoccupation majeure. Vous connaissez le grand débat actuel qui porte sur la baisse des gains de productivité au sein de l'OCDE. En France, il s'agit d'une priorité puisque nous sommes passés sous la barre des 1 % depuis deux ans. Or il est manifeste que les technologies d'intelligence artificielle sont capables d'appréhender des environnements d'incertitude élevée et de créer de la productivité dans bien des domaines : recherches documentaires, recherches juridiques et jurisprudentielles, transports de personnes... Je pense par exemple aux clercs de notaires ou aux conducteurs de véhicules. Il est par conséquent essentiel d'aider le plus largement possible à la diffusion de ces techniques, dans la mesure où des enjeux de compétitivité entre les nations s'expriment à cet égard. Il faut à la fois faire de la recherche fondamentale et de la recherche appliquée.

Il convient de mettre en place des modèles d'apprentissage et de diffusion de la science les plus pertinents possibles. Aussi bien en termes épistémologiques que d'enseignement, le numérique implique des nouvelles formes de transmission des savoirs.

Nous avons tout d'abord besoin d'interdisciplinarité. Je ne reviendrai pas sur les notions et les enjeux de pluridisciplinarité très largement évoqués précédemment, mais j'évoquerai une nouvelle fois la thématique des mondes ouverts.

Les mondes ouverts interviennent aujourd'hui dans une dynamique qui ne peut plus être ignorée, avec une efficacité bien supérieure au système de recherche verticale. Je suis très sensible au fait que nous disposions en France d'institutions telles que l'École Normale Supérieure (ENS) qui, en termes de ratios sur le nombre de prix Nobel, est en tête de tous les classements. Pour autant, il s'agit d'une école très élitiste donc limitée en

nombre d'élèves. Nous pouvons saluer les travaux essentiels réalisés au sein de l'ENS dans le domaine de l'intelligence artificielle, mais ces travaux ne me semblent pas rendre compte des dynamiques du monde à venir. La plateforme d'*open source* *GitHub*, qui réunit à elle seule plus de seize millions de codeurs, produit trente-sept millions de projets informatiques, dont entre quarante et soixante mille projets dans le domaine de l'intelligence artificielle équivalant à plus de cent cinquante milliards de lignes de code. Or il n'y a aucune réflexion stratégique de l'enseignement supérieur et de la recherche pour massifier les dynamiques rencontrées sur ces plateformes et bâtir des ponts avec ces systèmes de science ouverte. Il faudrait systématiser au sein des institutions de recherche françaises et de l'enseignement supérieur des modèles reposant sur des systèmes d'*open source*.

Pour ma part, je suis proche de *La Paillasse*, laboratoire de sciences ouvertes dans le domaine du vivant, que j'accompagne de jour en jour pour l'aider à subsister. Les subventions dont il bénéficie sont assez limitées. En réalité, je constate qu'il existe une incompréhension absolument totale des institutions en charge des enjeux de recherche, dans un monde où la recherche devient de plus en plus futuriste alors que précédemment elle était incrémentale. En effet, la recherche du XX^e siècle s'effectuait en blouse blanche, pour inventer par exemple le moteur à explosion qui a sans cesse été amélioré par la suite. Puis est apparue au XXI^e siècle la personne ayant recueilli le plus de critiques positives dans toute l'histoire de l'automobile pour son véhicule Tesla, alors que sa société n'a que dix ans. Il ne faut pas se leurrer : les techniques de l'innovation mises en œuvre au sein de Tesla s'inspirent du monde du numérique, de *open source*, de la pluridisciplinarité et également de contributions externes.

Ces modes de fonctionnement sont très incompris d'éminentes personnalités qui dirigent les unités de recherche ou les universités les plus prestigieuses en France. Il s'agit d'une interpellation de ma part, car je crois important de s'intéresser à ces domaines. En effet, même si l'expertise mathématique continuera d'être essentiellement académique, une relation étroite entre algorithmie et code est mise en œuvre pour faire fonctionner les intelligences artificielles. Autant je peux comprendre que l'algorithmie nécessite des chaires académiques et des laboratoires de recherche, autant la partie code est davantage profane. Tout le monde peut contribuer à innover dans le domaine du code.

En outre, il m'apparaît indispensable de nous intéresser à ces systèmes d'innovation ouverte - et je tiens ces propos avec ma casquette européenne - dans la mesure où le niveau d'investissement de la Chine et des États-Unis dans le domaine de l'intelligence artificielle est devenu considérable. La Chine serait même passée devant les États-Unis, avec la recherche conduite par Baidu qui représenterait à elle seule le tiers du total des dépenses de recherche en intelligence artificielle des États-Unis. De ce fait, nous ne pourrions plus entrer en compétition sur la quantité, mais sur le modèle de science que nous mettrons en œuvre, ce sur quoi nous devons

concentrer nos efforts. Or selon un rapport de 2016 de l'OCDE, la France est le pays dans lequel la collaboration entre l'entreprise et l'université est la plus faible. Certes nous pouvons, comme la représentante de Facebook l'a fait, nous féliciter de l'existence des CIFRE et du CIR, mais la transdisciplinarité doit également être encouragée.

J'achèverai mon propos sur la nécessité de créer des ponts entre le public et le privé et entre le primaire, le secondaire, le supérieur et la recherche. Je ne crois pas au futur de la recherche académique au sens strict, car les contributions extérieures seront essentielles à la réussite. Une personne autodidacte ayant appris à coder depuis l'école primaire pourra fournir des apports conséquents dans ce domaine. Il faut une prise de conscience politique sur ce point sous peine d'être relégué à un rang secondaire.

Mme Dominique Gillot, rapporteure. – Vous rejoignez les premiers intervenants sur le rôle de l'éducation et la formation des esprits à une plus grande agilité pour permettre éventuellement les formations autodidactes. Vous avez, les uns et les autres, apporté des informations qui vont susciter le débat.

8. Débat

M. Igor Carron, entrepreneur, fondateur du Paris Machine Learning meetup. - J'ai entendu les représentants d'Inria ainsi que les intervenants de Google et Facebook, nous tenir des propos très intéressants. En revanche une chose n'a pas été dite. *Scikit-Learn*, bibliothèque logicielle libre dédiée à l'apprentissage automatique, est un outil de création française, utilisé par environ deux cent mille personnes, des ingénieurs du monde entier. Cet outil, utilisant l'*open source* et en partie financée par Inria, n'a pas été mentionné. Lui, tout comme un certain nombre d'autres utilisés par nos start-up et notre tissu économique, sont absents des discussions. Je parle à la représentation nationale sur ce point. *Scikit-Learn* est réalisé en grande majorité en France, et est utilisé par une importante communauté à l'heure actuelle. D'ailleurs cet outil pourrait être accessible à des élèves dès la seconde. Il me semblait par conséquent important de rappeler ici que des projets existent, qu'ils sont sous les radars, y compris de personnes qui parlent d'intelligence artificielle avec autorité, et qu'ils éprouvent de grandes difficultés à survivre.

Mme Dominique Gillot, rapporteure. – Le cas de *Scikit-Learn* me rappelle celui d'*OpenClassrooms*, qui lui aussi est invisible.

M. Igor Carron, entrepreneur, fondateur du Paris Machine Learning meetup. – Oui, mais le modèle d'*OpenClassrooms* se rapproche plus de celui des MOOC (*massive open online courses* ou formation en ligne ouverte à tous).

M. Bertrand Braunschweig, directeur du Centre Inria de Saclay. - Merci à Igor Carron d'évoquer *Scikit-Learn*, dont une partie est développée à Inria de Saclay. La communauté est bien consciente de son importance et se montre soucieuse de son maintien. Je voulais vous rassurer sur le fait que *Scikit-Learn* est plutôt en bonne santé aujourd'hui.

M. Pavlos Moraitis, directeur du laboratoire d'informatique de l'Université Paris-Descartes. - Les financements public-privé nécessitent un meilleur développement en France. Les recherches universitaires doivent davantage bénéficier de financements par les grandes entreprises. À titre d'exemple, l'un de mes collègues d'Harvard m'a récemment raconté qu'il lui était tout à fait possible d'appeler un représentant de Google au téléphone, pour obtenir dès le lendemain un financement sur un projet.

M. Olivier Esper, responsable des affaires publiques de Google France. - Je tiens à signaler que nous offrons des dispositifs similaires en Europe.

M. Olivier Guillaume, Président d'O² Quant. - Nous constatons que l'intelligence artificielle est uniquement un outil d'aide à la productivité en participant à l'augmentation cognitive de l'homme. En France, nous avons les meilleurs cerveaux en recherche. Comment pourrait-on favoriser l'émergence de l'intelligence artificielle dans l'innovation française et européenne et permettre l'amélioration de nos sociétés ?

M. Jean-Daniel Kant, maître de conférences à l'Université Pierre-Marie-Curie Paris-VI. - J'ai trouvé l'analyse de Dominique Cardon passionnante et je voudrais conseiller de regarder l'excellente mais très déprimante série *Black Mirror*, qui met en scène des projections très réalistes sur une intelligence artificielle mal gérée. Mon objectif n'est pas de produire des intelligences artificielles autonomes, mais je tiens à souligner que ces intelligences existent déjà de fait.

M. Jean Ponce, professeur et directeur du département d'informatique de l'École normale supérieure (ENS). - Pas vraiment...

M. Jean-Daniel Kant, maître de conférences à l'Université Pierre-Marie-Curie Paris-VI. - Les *Google Cars* en font partie. Il y a aussi le cas de conseils d'administration où siègent des intelligences artificielles. Je n'ose pas imaginer les projets de recherche et de défense que nous ne connaissons pas, et dans lesquels l'intelligence artificielle autonome est également présente. Certes, il ne s'agit pas de la meilleure façon de répondre positivement aux peurs, mais je pense qu'il y a un certain nombre de gens qui souhaitent, pour une série de raisons, développer l'intelligence artificielle autonome. Il ne faut, par conséquent, pas en nier l'existence.

M. Patrick Albert, PDG de SuccessionWeb. - J'ai une question qui me semble fondamentale à l'intention des parlementaires ici présents. Quels sont les objectifs que nous pouvons nous donner en tant que pays ? Alors que le titre de la session est « *dimensions stratégiques en matière d'intelligence artificielle* », nous voyons que trois intervenants représentent des

multinationales basées aux États-Unis. Certes nous pouvons nous réjouir qu'ils viennent en France travailler avec nous, mais il est tout de même inquiétant que ce soient eux qui nous conseillent, surtout à l'heure où il s'agit de rendre à nouveau grande l'Amérique. Les brevets déposés sont largement américains. Je pense par conséquent que nous devons travailler à définir nos objectifs nationaux.

Mme Dominique Gillot, rapporteure. - Merci pour cette question qui resitue le sujet en son cœur.

M. Yves Demazeau, président de l'association française pour l'intelligence artificielle (AFIA). - Je souhaite signaler que nous organiserons lundi prochain avec le MEDEF une journée « *Intelligence artificielle et entreprises de France* », au cours de laquelle seront présents des groupes tels que Renault, Veolia ou Dassault pour évoquer l'impact de l'intelligence artificielle sur leurs activités.

M. Jean-Claude Heudin, directeur de l'Institut de l'Internet et du Multimédia. - Nous disposons aujourd'hui d'outils de financement de la recherche mais je tiens à insister sur la complexité des dossiers, qui représente un vrai frein à la compétitivité. On passe plus de temps à monter les dossiers et dans les contrôles que dans le travail de recherche lui-même. Aujourd'hui les start-up ne peuvent pas se permettre de dépenser autant d'énergie à travailler sur les dossiers. Le financement obtenu avec des mois de retard représente un véritable problème.

Mme Dominique Gillot, rapporteure. - Ces travers ont été largement commentés mais on assiste à une simplification des procédures, le futur livre blanc de l'enseignement supérieur et de la recherche l'illustrera.

Mme Amal Taleb, vice-présidente du Conseil National du Numérique. - Je travaille pour SAP et je suis membre du Conseil National du Numérique. Je souhaite partager une question sous forme de constat avec vous, sur laquelle nous devrions réfléchir collectivement. L'intelligence artificielle doit investir le tissu industriel classique. La France a la chance de posséder un beau tissu industriel. L'Allemagne a récemment fait le même constat lors de la conférence franco-allemande. Pour autant, la représentation nationale ne pourra faire l'économie de la manière dont l'intelligence artificielle va intervenir directement sur les *process* de production *B to B* me semble-t-il, en mettant les entreprises, les grandes comme les petites, dans la discussion. Il faut de l'intelligence collective et garder à l'esprit le fait que nous sommes engagés dans une industrie du futur, dans laquelle l'intelligence artificielle devra être prise en compte de façon importante. Certes nous avons la culture *B to C* de l'usage, des consommateurs et des usagers, mais il faudra de manière beaucoup plus globale introduire en mode *pull* l'intelligence artificielle comme un moyen de gagner des parts de marché.

Mme Dominique Gillot, rapporteure. - Différentes stratégies ont été mises au point par des comités qui ont réfléchi longuement aux points

que vous évoquez. Sur la stratégie nationale de la recherche, les mesures prises il y a un an sont déjà rediscutées eu égard à la rapidité avec laquelle les sujets et les questions prégnantes émergent. Effectivement, nous devons mener un débat entre les représentants de la recherche fondamentale académique et cette poussée très forte des jeunes *start-uppers*, qui réinvestissent des connaissances acquises il y a parfois longtemps et mises à jour de manière autodidacte. De jeunes diplômés tentent également l'expérience de la start-up sans choisir la recherche fondamentale. Nous ne pouvons pas non plus exiger que les choses s'effectuent de manière instantanée. L'ensemble de vos remarques sont essentielles et seront consignées dans notre rapport. Un bouleversement devra manifestement être opéré dans notre société sur la transformation de la connaissance en innovation, tout en n'omettant pas de tenir compte des impératifs de la décision politique, qui doit être prudente. Nous sommes en France.

Mme Amal Taleb, vice-présidente du Conseil National du Numérique. - Effectivement vous avez parfaitement raison, il nous faut résoudre ce hiatus. L'un des moyens est d'investir le tissu industriel beaucoup plus classique. Il se trouve que nous avons beaucoup de chance, puisque que nous avons des logiques et des réflexions similaires à celles de l'Allemagne. De manière différente, ce pays intègre les entreprises dans son tissu. Il nous faut impérativement inclure les grands groupes, les PME plus classiques et les start-up afin de créer cet écosystème qui permettra de développer l'intelligence artificielle et sans doute, de faire jouer les intelligences collectives. Ces associations pourraient également nous permettre de dépasser les blocages financiers.

M. Gilles Dowek, directeur de recherche Inria et professeur attaché à l'ENS de Paris-Saclay. - La difficulté à trouver des financements ne vaut pas seulement pour le secteur privé, mais également pour des organismes publics tels que le CEA.

M. Jean Ponce, professeur et directeur du département d'informatique de l'École normale supérieure (ENS). - Nous n'avons pas, à l'ENS ou à Inria, de réelles difficultés à trouver des financements de la part d'industriels car nous en rencontrons chaque semaine. Nous sommes privilégiés à ce niveau. Les réseaux de chercheurs (RC) sont aussi un bel outil à cet égard. En revanche, j'explique à ces partenaires que nous sommes une petite équipe en recherche fondamentale, qui ne dispose pas du temps et des compétences nécessaires au travail d'encadrement. Le statut de chercheur n'est pas fait pour cela, et pour ma part je n'aurais pas le temps d'encadrer un trop grand nombre de post-doctorants, de l'ordre de la dizaine par exemple.

Mme Dominique Gillot, rapporteure. - Nous avons évoqué le plan de carrière d'un chercheur et les différentes étapes à passer, qui nécessitent des allers-retours. Votre questionnement a été abordé à travers la stratégie nationale de l'enseignement supérieur : comment valoriser les parcours

itératifs d'un chercheur dans le privé, le public, la recherche et la formation ? Aujourd'hui nous n'avons pas les bons outils.

M. Jean Ponce, professeur et directeur du département d'informatique de l'École normale supérieure (ENS). - Aux États-Unis, des collaborateurs ont pour mission de travailler avec les industriels et de faire le pont avec les chercheurs.

M. Fabien Moutarde, professeur au centre de robotique de Mines ParisTech. - Il faut faire évoluer les statuts des chercheurs pour permettre davantage d'allers-retours entre la recherche privée et la recherche publique. J'ai commencé ma carrière chez Alcatel, mais j'ai eu du mal à revenir dans le public.

Sur la partie stratégique, mon sentiment est qu'il est nécessaire de pousser au maximum, par des plans de financement de projets de recherche, la robotisation et l'automatisation de tout ce qui peut l'être. Il s'agit en effet d'un enjeu économique et stratégique pour la France et pour l'Europe. Or une censure a existé sur le sujet par crainte du chômage. Pourtant, l'absence d'automatisation aboutira malgré tout à la perte des emplois que l'on souhaite protéger. Des complémentarités saines peuvent être trouvées entre la robotique et l'humain, pour confier aux hommes les tâches ayant davantage de valeur ajoutée. On peut aussi aller vers un revenu citoyen et conduire à des relocalisations.

Mme Dominique Gillot, rapporteure. - Les sciences dures et les sciences sociales doivent davantage travailler ensemble. Il semble en effet souhaitable de supprimer les tâches répétitives et difficiles, pouvant entraîner à la longue un handicap. Il est vrai que dans le débat actuel de la société, cette dimension du bénéfique pour les personnes dont les métiers vont être supprimés n'est jamais abordée. C'est pourquoi le sujet de l'éducation est essentiel.

M. Igor Carron, Paris Machine Learning Meetup. - En France nous avons un problème avec les données. Alors que certains organismes parviennent à accumuler un grand nombre de données, ils sont immédiatement confrontés aux limites de leur utilisation. Au ministère de l'Éducation, il ne semble pas que des recherches soient menées sur le développement et l'éducation afin d'aboutir à une meilleure efficacité, aucun centre ne collecte les données pour améliorer les politiques éducatives à ma connaissance. Facebook et Google n'étaient que des start-up il y a encore dix-quinze ans, mais se sont développées grâce à l'intelligence artificielle. Pour ma part, j'estime hautement souhaitable que les données de l'Éducation Nationale puissent être utilisées par les chercheurs afin de faire progresser la science de l'éducation. Il nous manque un centre d'expertise à ce sujet.

Mme Laurence Devillers, professeur à l'Université Paris-Sorbonne/LMSI-CNRS. - Je souhaite intervenir sur la problématique de l'éthique, au sujet de laquelle les intervenants de Facebook et Google ont mentionné qu'une réflexion était en cours au sein d'un partenariat entre

entreprises. Existe-t-il des moyens de mener cette réflexion en collaboration avec les chercheurs ? Comment allez-vous vous ouvrir alors qu'on constate une certaine opacité ? Les impacts d'une telle réflexion porteront notamment sur l'éducation et la protection des données.

M. François Taddei, directeur du Centre de recherche interdisciplinaire. – Aujourd'hui, les traces numériques des apprenants ne sont pas disponibles pour faire progresser l'ensemble du système. Pourtant, ces données pourraient être un miroir pour l'enseignant, le chef d'établissement et les parents. Si nous pouvions passer des *data* à l'information, nous progresserions grandement. T. S. Eliot se demandait où est passée la sagesse que nous avons perdue dans la connaissance et où est passée la connaissance que nous avons perdue dans l'information... Nous avons besoin de créer des centres de recherche pour avancer sur ces sujets. La CNIL et Inria dialoguent à ce sujet. Il faut mettre en place des plateformes sécurisées en termes de vie privée, mais avec une vraie valeur collective. Le *big data* pourrait par exemple proposer des solutions pour résoudre les difficultés d'apprentissage. Inria s'est beaucoup rapprochée de l'INSERM pour travailler dans le domaine de la recherche en éducation, mais nous ne disposons d'aucune possibilité d'analyser les *data*. Nous savons qu'on consacre 2,24 % de la richesse nationale à la R&D mais nous ne savons pas combien nous consacrons à l'éducation.

Mme Dominique Gillot, rapporteure. – Vous soulevez de vraies questions. La recherche en éducation doit se faire à l'université.

M. François Taddei, directeur du Centre de recherche interdisciplinaire. – Cependant, la capacité des enseignants sur l'intelligence artificielle et le numérique est très limitée.

Mme Dominique Gillot, rapporteure. – Elle est très réduite, en effet. Il faut agir sur les structures de formation des enseignants, en particulier les écoles supérieures du professorat et de l'éducation (ESPE), qui ont remplacé les instituts universitaires de formation des maîtres (IUFM). La définition des maquettes est rigidifiée par les anciennes pratiques. J'ai assisté à un débat absolument incroyable au cours duquel on expliquait qu'un poste réservé à un mathématicien ne pouvait pas être transformé en un poste d'enseignant en informatique. Les carrières sont, à l'évidence, beaucoup trop rigides. À l'Université de Cergy-Pontoise, une recherche est en cours sur l'utilisation des outils de l'intelligence artificielle pour observer comment les enfants fixent leur attention. La difficulté actuelle tient à la résistance des parents à ce que leurs enfants soient l'objet de telles expériences, car l'anonymisation n'est pas au rendez-vous. Il faut donc obtenir une réassurance en la matière.

M. François Taddei, directeur du Centre de recherche interdisciplinaire. – Il faut s'assurer que les progrès bénéficieront à chacun des acteurs, y compris aux premiers intéressés. Si les parents ont l'impression que les progrès bénéficient à certains seulement, nous n'y arriverons pas.

M. Gilles Dowek, directeur de recherche Inria et professeur attaché à l'ENS de Paris-Saclay. – Aujourd'hui, le ministère de l'Éducation nationale ne comprend rien aux enjeux qu'il y a à former les jeunes sur l'informatique et l'intelligence artificielle. Je donnerai un exemple sur ce point. Il est coutume de dire que pour enseigner un sujet, quatre éléments sont nécessaires : un élève, un enseignant, du temps à consacrer à cette activité, et un programme. Aujourd'hui, le Conseil supérieur des programmes a fait un bon travail pour élaborer des contenus à enseigner en informatique au collège, mais il manque des enseignants. Or le ministère de l'Éducation nationale ne comprend pas la nécessité d'avoir des enseignants en informatique. Le raisonnement au sein de ce ministère est d'affirmer que l'informatique étant une discipline relativement facile, les jeunes possèdent la matière dans leur ADN. De ce fait, tant que nous resterons arc-boutés sur ce type de conceptions, nous ne pourrons pas avoir un public informé de ce que les algorithmes sont censés faire.

M. Jean-Gabriel Ganascia, professeur à l'Université Pierre-et-Marie-Curie-Paris VI. – Dans tous les métiers, le numérique est central. Il faut par conséquent lancer des formations mixtes, car c'est nécessaire pour le secondaire comme pour le supérieur. Pour notre part, nous avons élaboré une formation lettres-informatique, car il faut nous préparer à de nouveaux métiers. Nous avons parlé du travail et des risques que la robotisation faisait courir à certains métiers, mais il est possible de compenser les pertes d'emplois par de nouveaux métiers. Pour préparer les personnes à ces métiers, la formation est tout à fait essentielle. À cela, il faut ajouter que la formation dont il s'agit n'est pas nécessairement initiale. La formation tout au long de la vie doit être développée, en transformant les institutions supérieures et les écoles pour parvenir à ce résultat. Il faut également revenir sur le travail, pour lequel la comptabilisation en « temps passé à l'établi » n'a plus aucun sens. Le temps de présence au bureau c'est obsolète. Il est indispensable d'ajouter le temps de formation au temps de travail, étant observé que l'idée de trente-cinq heures passées au travail est tout à fait désuète. Je pense que nous aurons du mal à faire passer ces idées.

Pour rejoindre les propos de Gilles Dowek sur l'administration de l'Éducation nationale, je confirme que le ministère a du mal à opérer ces transformations. J'avais moi-même un projet de sciences participatives des enseignants dans les domaines littéraires, mais ce projet a été mis sur la touche faute de pouvoir être mis en contact direct avec des enseignants par un directeur académique au numérique, qui était aussi peu compétent qu'indisponible.

Enfin, sur les questions d'éthique et d'intelligence artificielle, nous organiserons avec l'AFIA le 5 juillet 2017 une journée qui leur est consacrée.

Mme Dominique Gillot, rapporteure. – Juste un petit rappel sur le compte personnel d'activité (CPA), preuve que les parlementaires ont malgré tout fait avancer les choses en matière de formation. Le CPA a été mis en place au 1^{er} janvier 2017 pour associer formation et travail. Nous voulons

également confier à l'université la formation tout au long de la vie. La loi doit être appliquée, mais elle met du temps pour ce faire.

M. Gilles Babinet, entrepreneur, *digital champion* auprès de la Commission européenne. – Nous avons des approches très holistiques, comme en témoignent les exemples de nos réflexions sur l'éthique ou sur l'éducation. Cela nous ramène à l'un des problèmes dont souffre notre pays, qui est celui de la parole politique, handicapée par l'incapacité du corps politique à se projeter dans un futur humaniste, et qui respecte nos valeurs. Il faut tracer ce futur. Trop souvent, les débats tiennent à des enjeux du XX^e siècle. Or, la révolution digitale ne peut plus nous faire penser que nos sociétés fonctionneront comme au siècle dernier.

Mme Dominique Gillot, rapporteure. – L'OPECST a été créé dans ce but entre autres. Je voudrais que Google réponde aux questions qui ont été posées concernant le partenariat sur l'intelligence artificielle. Pourra-t-on faire des choses ensemble ?

M. Olivier Esper, responsable des affaires publiques de Google France. – Je réponds positivement à la question de Laurence Devillers. Les uns et les autres, avons bien souligné nos modèles ouverts de recherche, qui ont évolué dans ce sens dans la période récente. Le partenariat sur l'intelligence artificielle a bien évidemment pour objet de regrouper la diversité. Son conseil d'administration n'est pas encore finalisé mais des Européens dont Yann LeCun, français, en font déjà partie. L'idée est de mettre en œuvre une réflexion collective.

Mme Dominique Gillot, rapporteure. – L'ouverture est une bonne chose mais il s'agit également de travailler avec des partenaires non liés à vous par contrat.

M. Olivier Esper, responsable des affaires publiques de Google France. – Sur le site du partenariat sur l'intelligence artificielle figure déjà la possibilité de manifester son intérêt à participer à la réflexion.

Mme Delphine Reyre, directrice Europe des affaires publiques de Facebook. – Nous nous engageons auprès de vous à faire part de l'intérêt fort exprimé à être parties à la réflexion. Nous pourrions ensuite imaginer les formes de cette discussion.

Mme Laurence Devillers, professeur à l'Université Paris-Sorbonne/LMSI-CNRS. – Je réfléchis avec plusieurs collègues sur la robotique et l'éthique dans le cadre de la CERNA. Nous avons donc des sujets de réflexion qui pourraient intéresser les grands groupes. Nous vous transmettrons nos rapports, mais nous pourrions également imaginer autre chose dans une démarche participative avec d'autres groupes industriels. La réflexion sur l'intelligence artificielle doit être menée de manière collégiale en y intégrant les dimensions éthiques. Je suis impliquée dans la société savante IEEE, qui réunit des centaines de chercheurs du monde entier pour travailler sur des standards à mettre en place ensemble. Il est essentiel de ne

pas opérer de façon isolée sur des sujets tels que ceux-ci, qui concernent notre futur. C'est pourquoi nos gouvernements devraient s'emparer de cette question.

IV. TROISIÈME TABLE RONDE, PRÉSIDIÉE PAR MME DOMINIQUE GILLOT, RAPPORTEURE : QUESTIONS POLITIQUES, SOCIÉTALES ET ÉCONOMIQUES LIÉES À L'IRRUPTION DES TECHNOLOGIES DE L'INTELLIGENCE ARTIFICIELLE

Mme Dominique Gillot, rapporteure. – Nos deux tables rondes de ce matin, qui ont donné lieu à des échanges intéressants, ont permis de cerner ce que sont les systèmes d'intelligence artificielle, dans leur diversité.

Cette troisième table ronde aborde les questions politiques, sociétales et économiques liées à l'irruption des technologies de l'intelligence artificielle. Axelle Lemaire, secrétaire d'État chargée du numérique et de l'innovation, nous rejoindra à seize heures.

Notre calendrier est contraint mais nous souhaitons tous aller au bout de nos réflexions avant les prochaines échéances électorales.

Monsieur Henri Verdier, vous êtes directeur interministériel du numérique et du système d'information de l'État, adjoint à la secrétaire générale pour la modernisation de l'action publique et administrateur général des données. Vous avez dirigé Etalab, le service chargé de l'ouverture des données publiques, grâce auquel vous avez modernisé le portail d'*open data* français. Votre regard sur les questions politiques, sociétales et économiques liées à l'irruption des technologies d'intelligence artificielle sera utile pour ouvrir cette table ronde.

1. M. Henri Verdier, directeur interministériel du numérique

L'intelligence artificielle est une vieille histoire qui se renouvelle actuellement à une allure incroyable. Se sont télescopées de manière imprévue une science, une masse de données et une informatique, venue en particulier des jeux vidéo, avec des investissements considérables de la part des acteurs qui s'en servent : Google, IBM, Facebook... Il s'agit d'investissements qui atteignent des niveaux inédits.

Ces acteurs s'inscrivent dans des stratégies post-numériques ouvertes et atteignent d'amples performances. Nous avons récemment appris, par exemple, que l'intelligence artificielle reconnaissait mieux les images que l'œil et le cerveau humain !

Certains orateurs ont assimilé cet enjeu à celui qu'a représenté la course à l'arme nucléaire. La métaphore a ses limites : il ne s'agit pas forcément de fabriquer des intelligences artificielles mais il s'agit d'en avoir

la maîtrise. Certains pays maîtriseront cette technologie, d'autres non, ce qui suscitera un clivage géopolitique.

Pour l'État, le premier enjeu est que notre pays ne devienne pas, comme le craint **Mme Catherine Morin-Desailly**, une colonie numérique des États-Unis. En ce qui concerne le *big data*, la France n'a que deux ou trois ans de retard : il est donc encore temps de démarrer. La bataille n'est pas perdue. Notre pays peut compter sur une très grande école de mathématiques, de très grandes écoles d'ingénieur, une culture de l'entrepreneuriat, des industriels de renom, un État centralisé qui pourrait devenir un client. Il faut s'y mettre, avec une politique industrielle moderne. Je relève des initiatives très positives : le travail de l'OPECST sur l'intelligence artificielle, le plan gouvernemental pour l'intelligence artificielle, etc.

Des questions de sécurité se poseront. Je ne suis pas informaticien de formation mais biologiste. J'ai une vague intuition des propriétés émergentes des systèmes. Dès qu'un système est complexe, son évolution est imprévisible. J'ai beaucoup milité en faveur de la publication des algorithmes. Cependant, l'intelligence artificielle fonctionne comme une boîte noire. Il ne suffit pas de divulguer le code source pour que chacun se fasse une opinion. La Silicon Valley met en place des *Kill Switch*, c'est-à-dire des dispositifs qui se désactivent très facilement.

En matière économique, on se prépare à de grandes mutations, de nouveaux cycles de destruction incroyables, de reconversion et de recombinaison de filières industrielles entières. Le marché du travail sera profondément transformé. Peut-être nous sera-t-il possible d'accompagner ces changements dès lors que cette révolution est prise à la racine.

L'État sera-t-il capable de s'approprier ces nouvelles technologies pour assurer l'efficacité, la simplicité et la justice du service public ? Pourquoi attendre un mois pour savoir si l'on a obtenu une bourse alors que l'on sait, lorsque l'on commande la moindre bricole sur Amazon, à quel moment elle se trouve mise dans la camionnette du facteur ?

Nous n'avons pas une très grande science du *design* des politiques publiques. Souvent, elles ont été construites de manière empirique.

J'ai, par ailleurs, été frappé de constater qu'il était très facile de faire des erreurs très idiotes avec des algorithmes. Par exemple, il est très simple de mettre au point un algorithme permettant de mener des contrôles fiscaux avec une efficacité de 90 %. Règle-t-on le problème ou fait-on courir des risques encore plus grands ? En effet, cet algorithme peut omettre une catégorie entière de fraudes, ce qui n'est pas souhaitable. Il faut donc bien être sûr, lorsque l'on construit un algorithme, qu'il est totalement adapté à la politique que l'on souhaite mettre en place. Or, il est souvent très difficile d'explicitier le but profond d'une politique publique.

On l'a vu avec le *big data*, il est assez aisé, faute d'une profonde réflexion, de faire faire des bêtises à des machines qui ne commettent

pourtant aucune erreur. Les États-Unis travaillent par exemple à la mise au point d'algorithmes pour aider les juges à prononcer des remises en liberté conditionnelle, mais on y a introduit des statistiques pleines de préjugés ethniques, ce qui conduit à une mauvaise prédictibilité. Les algorithmes donneront donc des résultats à la hauteur de l'éducation qui leur est fournie !

Nous avons le devoir de nous approprier ces nouvelles technologies, mais nous devons aussi nous demander où nous voulons aller. Cela nous amènera assez vite à la question de la « gouvernementalité algorithmique ». Il est utile de pouvoir prédire grâce aux algorithmes, mais cela comporte des risques. Notamment, celui de la confiscation du pouvoir par des experts. Il est important de faire entrer les vrais débats sur l'intelligence artificielle au Parlement et chez les citoyens. Les algorithmes apprennent et donnent des résultats mais ce sont des boîtes noires. Le besoin de transparence est très grand.

L'invention d'une redevabilité et d'une gouvernance de l'intelligence artificielle est un immense chantier. Si nous voulons rester une démocratie et utiliser à plein ce que proposent ces technologies, il nous faudra faire le lien entre projet politique et progrès technique.

Concernant le plan du Gouvernement pour l'intelligence artificielle, nous avons quelques éléments mais je préfère laisser la ministre en parler.

Mme Dominique Gillot, rapporteure. – Merci M. Verdier ! Vous présidez, Madame Marie-Claire Carrère-Gée, le Conseil d'orientation pour l'emploi (COE), qui vient tout juste de rendre public un rapport important sur les conséquences de l'automatisation et de la robotisation sur l'emploi. Je vous laisse la parole.

2. Mme Marie-Claire Carrère-Gée, présidente du Conseil d'orientation pour l'emploi (COE)

Le conseil d'orientation pour l'emploi est une instance indépendante, placée auprès du Premier ministre. Nous essayons d'élaborer des diagnostics partagés sur les questions d'emploi et de travail. C'est un domaine caractérisé par la pluralité des acteurs et dans lequel le consensus n'est pas toujours spontané.

L'ensemble des acteurs du marché du travail, comme l'ensemble de nos concitoyens, sont confrontés à l'évolution des nouvelles technologies, qui connaissent un extraordinaire développement et dont la vitesse de diffusion est très incertaine. Je précise que je n'ai pas de compétences en intelligence artificielle.

Notre rôle est d'anticiper, autant que faire se peut, les conséquences du progrès technologique en cours sur l'emploi, dans un contexte où le débat

public est marqué par des études donnant au pourcentage près le nombre d'emplois menacés et conduisant à un climat très anxiogène. Certes, la crainte du chômage est un grand classique à chaque vague d'innovation technologique. Keynes lui-même avait prédit un chômage massif lié au progrès technique. Pourtant, l'histoire montre que, depuis toujours, le progrès technologique s'est accompagné de créations d'emplois.

Malgré cela, des études prédisent une catastrophe à un horizon proche. Je peux citer entre autres l'étude de l'Université d'Oxford conduite par Carl Benedikt Frey et Michael A. Osborne qui parle de 47 % des emplois, celle du centre de réflexion Bruegel avec une part des emplois concernés de 52 % et, enfin, celle du cabinet de conseil Roland Berger, spécifique à la France, avec une part de 42 %. D'autres études sont moins pessimistes, avec un chiffre de 9 %, comme le rapport de l'OCDE réalisé par des chercheurs de l'Université de Mannheim ou, encore, la note que Nicolas Le Ru a réalisée pour France Stratégie en 2016.

Les analyses existantes sont partielles. Elles s'intéressent uniquement aux destructions brutes d'emplois, et ne font pas la balance entre les destructions et les créations. Elles ne s'attachent ni à la structure de l'emploi ni à sa localisation. Parfois même, elles sont biaisées en ce qu'elles considèrent que la technologie bouscule des métiers dans leur ensemble alors que, de notre point de vue, elle bouleverse d'abord des tâches. Au sein d'un métier, certaines tâches peuvent rester très difficiles, voire impossibles, à remplacer par des technologies. Toutes les personnes qui exercent un métier ne font par ailleurs pas la même chose tous les jours. Le rapport de l'OCDE s'intéresse aux tâches, ce qui est plus pertinent. Les emplois évoluent, y compris dans leur contenu.

Dans notre rapport, nous avons voulu aborder l'ensemble des dimensions - créations, destructions, transformations d'emploi, structures et localisation de l'emploi - et élargir l'approche, non seulement aux emplois susceptibles d'être détruits, mais aussi à ceux qui peuvent évoluer dans leur contenu. Ce constat doit conduire à définir des politiques de formation adaptées. Les études rétrospectives sur la composition de l'emploi montrent d'ailleurs que de nombreux emplois, en France, ont profondément évolué depuis quinze ans dans leur contenu en lien avec la diffusion des nouvelles technologies, quel que soit le niveau de qualification. L'utilisation de technologies numériques est associée à une complexification et une hybridation des compétences requises. Avec l'évolution numérique, un certain nombre de métiers se sont même scindés en plusieurs.

Dans ce cadre, nous avons réalisé une étude quantitative pour préciser le diagnostic, pour savoir quels seraient les métiers concernés. Nous avons émis l'hypothèse que la vulnérabilité des emplois dépend de trois facteurs principaux : l'avantage comparatif de l'homme par rapport aux machines ; les goulets d'étranglement de la frontière technologique ; enfin, la rentabilité économique. Difficile à introduire de façon robuste, ce dernier

facteur n'est pas pris en compte dans notre étude, pas plus d'ailleurs que dans les autres études existantes.

Nous avons utilisé l'enquête sur les Conditions de travail de la DARES. Nous avons isolé, selon les types de métier, jusqu'à une vingtaine de questions. Nous avons essayé d'apprécier le niveau de vulnérabilité des emplois au regard de leurs conditions concrètes de réalisation. Nous nous sommes refusé à donner un chiffre précis car il n'existe aucun seuil à partir duquel un emploi serait automatiquement automatisé. Nous avons préféré retenir des ordres de grandeur.

Nous sommes parvenus à trois enseignements principaux. Tout d'abord, moins de 10 % des emplois cumulent des vulnérabilités telles que, dans un contexte d'automatisation, leur maintien serait menacé. Nous avons retenu un seuil d'automatisation de 0,7 : à noter que les résultats auraient potentiellement été un peu différents, avec un seuil de 0,65 par exemple. Les métiers les plus vulnérables sont essentiellement peu ou pas qualifiés, nous en avons listé une vingtaine : agents d'entretien, ouvriers qualifiés ou non, manutentionnaires, jardiniers, maraîchers, caissiers, employés de service...

Ensuite, la moitié des emplois seront profondément, à plus ou moins brève échéance, transformés dans leur contenu, quel que soit le niveau d'automatisation. Cela touche les moins qualifiés mais aussi des emplois plus qualifiés, dans l'industrie mais aussi dans les services.

Enfin, l'étude quantitative n'était pas le seul sujet de notre rapport. Nous avons pris acte du fait que l'histoire économique montre que l'emploi a continué à augmenter, y compris au cours des vingt dernières années. L'enjeu est la gestion de la période de transition. Nous avons étudié les questions de création d'emplois : ceux induits directement par l'innovation technologique mais aussi ceux indirectement créés. Même une simple innovation de procédé, si elle fait gagner des parts de marchés à l'international, peut créer de l'emploi. Il y a aussi des mécanismes de compensation : l'évolution de la demande crée des emplois induits.

Une politique publique d'accompagnement doit viser à diminuer le nombre d'emplois détruits et faire en sorte que les mécanismes de compensation se produisent le plus tôt possible et jouent à plein pour éviter les décalages entre destructions et créations d'emplois. Le progrès technologique n'est pas une option que l'on pourrait ou non choisir ; on ne peut l'éviter et la vitesse d'adaptation est essentielle. Les acteurs économiques et les pouvoirs publics ont leur mot à dire en termes de choix éthiques, sociaux, économiques, fiscaux. Je vous renvoie aux conclusions de notre rapport. Si la première vague du numérique a contribué à délocaliser des emplois existants, voire des fonctions entières dans les entreprises, la vague en cours pourrait, *via* l'automatisation, favoriser des relocalisations d'emplois. C'est un aspect qui doit être pris en compte.

Mme Dominique Gillot, rapporteure. - Chirurgien neurologue, ancien élève de Science Po, de HEC, de l'ENA, écrivain, auteur d'ouvrages à

succès, chroniqueur pour plusieurs journaux, vous êtes aussi, monsieur Alexandre, le fondateur du site Doctissimo, très consulté, et vous dirigez aujourd'hui l'entreprise DNA vision. L'intelligence artificielle est un sujet qui vous passionne. Nous vous écoutons.

3. M. Laurent Alexandre, entrepreneur (DNA vision)

On parle souvent des dangers de l'intelligence artificielle forte, mais les vrais sujets sont la gestion, la gouvernance et la régulation d'une intelligence artificielle faible, sans conscience d'elle-même et pas directement dangereuse pour l'humanité.

Je souhaite insister sur la souveraineté numérique et sur le risque de devenir une colonie numérique des GAFAs (Google, Apple, Facebook et Amazon). Il existe un trou noir de valeur en faveur des géants du numérique qui maîtrisent l'intelligence artificielle. En la matière, nous sommes un pays du tiers-monde. Nous exportons des spécialistes, des cerveaux, et nous importons toute la journée de l'intelligence artificielle : soixante-cinq millions de Français importent de l'intelligence artificielle cent quatre-vingts fois par jour sur leur téléphone ! Autant dire que nous avons perdu notre souveraineté numérique.

Nous sommes à un tournant de l'humanité : tel est le slogan du patron d'Intel. Le patron de Google, Sergei Brin, affirme, quant à lui, qu'il va construire des machines qui raisonnent, pensent et feront les choses mieux que l'homme. Ça c'est de l'intelligence artificielle forte. Elon Musk, le patron de Tesla, expliquait le 2 juin dernier, dans un grand élan lyrique, qu'il était urgent de mettre des microprocesseurs dans nos cerveaux avant que l'intelligence artificielle ne nous transforme en animaux domestiques !

Cependant, nous ne sommes pas à la veille d'un HAL 9000 qui signerait la fin de l'humanité, comme dans « *2001, l'odyssée de l'espace* » ! Les enjeux actuels sont des enjeux sociétaux, des enjeux de souveraineté, des enjeux économiques, liés au fait que nous sommes nuls en matière d'intelligence artificielle.

Mais l'intelligence artificielle n'est pas seule. Les GAFAs et les BATX (Baidu, Alibaba, Tencent et Xiaomi) ont éclos. L'idéologie transhumaniste a fait pousser des pulsions démiurgiques, notamment dans les différentes Silicon Valleys ; ce que l'on appelle les NBIC (nanotechnologies, biotechnologies, informatique et sciences cognitives) arrivent à maturité ; la zone Asie-Pacifique est plus transgressive que nous.

Les NBIC sont des technologies exponentielles, donc très imprévisibles. L'intelligence artificielle va être gratuite en valeur relative par rapport à l'intelligence biologique. Cela a des conséquences majeures : quand un bien est gratuit, les substituts sont balayés et les complémentaires sont

renchérés. Or, le travail peu qualifié est substituable, et le travail qualifié, complémentaire. La micro-économie rend ce verdict incontournable.

Si je parle de pulsions démiurgiques, c'est que la zone Asie-Pacifique se croit tout permis sur le plan biotechnologique, mais aussi sur le plan de l'intelligence artificielle. Une bonne gouvernance et une régulation internationale seront nécessaires.

Autre phénomène, la « platformisation » : l'intelligence artificielle sort au robinet des grandes plateformes pour des raisons techniques, car il faut beaucoup de données pour la développer. Une étude récente a montré qu'un mauvais algorithme avec beaucoup de données est supérieur à un bon algorithme avec peu de données. Or les plateformes ont des milliards d'utilisateurs, il n'y a aucun équivalent en Europe, nous n'avons pas de base industrielle comparable. Bref, nous sommes tous les idiots utiles de l'intelligence artificielle, pour paraphraser Lénine qui disait que les bourgeois de gauche sont les idiots utiles de la Révolution. En effet, nous mettons chaque jour dix milliards de photos gratuitement sur la nébuleuse Facebook, Instagram, Messenger et WhatsApp afin d'améliorer l'efficacité des intelligences artificielles des plateformes.

Les technologies NBIC se développent, mais pas toutes au même rythme exponentiel : la biologie et la robotique, c'est lent. La robotique ne fera pas beaucoup bouger le marché du travail. L'intelligence artificielle sans robot va croître exponentiellement et modifiera radicalement des aspects importants du travail. En revanche, l'électronique et l'intelligence artificielle sans support mécanique et biologique vont croître beaucoup plus vite. Le COE sous-estime ce poids croissant de l'intelligence artificielle.

Allons-nous devenir une colonie numérique des GAFA ? Au-delà de l'État, la société civile et l'ensemble du tissu économique doivent comprendre que si nous n'avons pas de plateformes à l'échelle des Chinois et des Américains, nous serons le Zimbabwe de 2080 ! Nous sommes en train de perdre nos ressorts économiques. Il n'y a pas en Europe une miette d'intelligence artificielle et de technologies NBIC, qui constituent pourtant la base fondamentale de création de valeur ! Nous avons une responsabilité historique.

Il s'est créé un trou noir de valeur en faveur des algorithmes et de l'intelligence artificielle, avec un changement radical dans la structure de valeur. Traditionnellement, on avait des *business* verticalisés. Nissan produit des voitures, Danone des yaourts, Gaumont du cinéma et Badoit de l'eau. Aujourd'hui, nous assistons à une horizontalisation des métiers. On voit arriver des conglomérats de l'intelligence artificielle, qui font plusieurs métiers. Baidu, le Google chinois, se lance dans la voiture autonome. L'intelligence artificielle produit cette transversalité. Les cinquante-cinq petits génies de WhatsApp ont créé en quatre ans 23 milliards de dollars de valeur, soit deux fois plus que les cent trente mille ouvriers de PSA depuis

un siècle. La paupérisation relative de notre population serait une certitude si nous ne prenions pas le tournant de l'intelligence artificielle !

L'intelligence artificielle induit un changement radical dans l'organisation de la fonction de production. On voit apparaître une part nouvelle de capital, l'intelligence artificielle, un déclin fort du travail et une montée en flèche de la valeur du travail très qualifié, qui est en réalité du capital, payé en actions et en stock-options. L'intelligence artificielle déstabilise le rapport relatif entre le travail et le capital, avec une nouvelle forme de capital à coût nul face au cerveau biologique. Cela ne veut pas dire que le travail va disparaître. En tant que schumpétérien de centre-gauche, je suis même persuadé que nous allons inventer de nouveaux métiers.

Quelles tâches pour quels hommes ? C'est un vrai sujet de réflexion. L'école de la République en France est extrêmement en retard, ce qui n'est pas le cas à Singapour, au Vietnam ou en Corée. La complémentarité neurone-transistor nous semble naturelle mais c'est en réalité un combat. En 2050, 100 % de ceux qui ne seront pas complémentaires de l'intelligence artificielle seront soit au chômage soit dans un emploi aidé ! Il est donc important de rénover progressivement l'école et la formation professionnelle. Ayant visité récemment des centres pour apprentis, je puis vous garantir que nous n'en prenons pas le chemin. Notre responsabilité politique est très importante, je ne suis pas sûr que nous nous en rendions bien compte. Nous vivons déjà des crises populistes, ça ne peut que s'aggraver.

Le défi posé par la complémentarité est le suivant : un robot coûte très cher, ce qui n'est pas le cas de l'intelligence artificielle, qui, de surcroît, progresse plus vite ! En outre, les robots remplacent plutôt des métiers peu qualifiés, alors que l'intelligence artificielle non fondée sur des robots peut remplacer des métiers très qualifiés. Je partage l'opinion de Yann LeCun, le patron d'une des deux branches de l'intelligence artificielle chez Facebook : l'intelligence artificielle aura dépassé les meilleurs radiologues - je me permets de le dire en tant que chirurgien, et ce n'est pas une vendetta corporatiste - avant 2030. Se pose donc un problème de reconversion qui dépasse celui des chauffeurs routiers et des emplois non qualifiés : l'intelligence artificielle rendra également nécessaire une réorientation de métiers traditionnels à haut contenu cognitif. D'ailleurs, le lien entre qualification des emplois et risque d'automatisation est beaucoup plus complexe qu'on ne l'imaginait il y a dix ans encore - mais on raisonnait en termes de robots au lieu de raisonner en termes d'intelligence artificielle !

Prenez la médecine personnalisée, qui traite chaque patient en fonction de ses caractéristiques génétiques : elle va tuer le médecin traditionnel - bien entendu, il ne s'agit que d'un raisonnement *ex ante* : les radiologues et les médecins rempliront de nouvelles fonctions. Le malade devient un ensemble de *data* et le médecin un *software*, ce qui n'empêche pas ce dernier d'être empathique ; le séquençage ADN représente trois milliards de bases ADN, vingt téraoctets de données brutes par malade ; chaque

individu représente deux millions de mutations par rapport au génome de référence – toutes ne sont pas promesses de maladie, bien sûr –, dont l'analyse exige des ACP, des analyses en composantes principales, à huit cents millions de facteurs. Dans un tel monde, il est clair que même le Dr House, en quinze minutes de consultation, déshabillage et rhabillage compris, sera impuissant. Nous allons donc assister à un transfert de pouvoir médical vers les algorithmes, la médecine étant désormais fondée sur un *big data* dont l'analyse par le cerveau biologique est exclue. Cela ne veut pas dire que le cerveau biologique n'a plus sa place dans cette équation ; mais cette place est profondément transformée. Cela n'est pas sans poser problème : aucun médecin n'étant capable d'analyser vingt téraoctets de données en quinze minutes, les décisions éthiques seront implicites dans l'algorithme d'intelligence artificielle.

Quelle réponse politique peut-on apporter à cette situation ? Je regrette qu'il n'y ait pas davantage de parlementaires présents cet après-midi. Le problème est réel : vous connaissez tous cette courbe qui fait froid dans le dos, celle de l'augmentation des revenus par décile en trois décennies de mondialisation. Tout le monde en a profité, sauf le petit blanc moyennement ou peu qualifié, c'est-à-dire les classes moyennes occidentales traditionnelles, qui ont vu leurs revenus stagner, voire, dans certains pays comme les États-Unis ou l'Allemagne, baisser de 1 à 2 % sur 30 ans. Ce qui est en jeu, en l'occurrence, ce n'est pas tant l'intelligence artificielle, mais, tout simplement, la confrontation entre l'Asie et l'Europe : un docteur en génomique à Bangalore coûte, à l'heure, trois fois moins cher qu'un ouvrier spécialisé qui fait les 35 heures en région parisienne. Cette situation ne durera pas ! Nous avons là l'une des explications à l'émergence du populisme et à la « *trump-brexitisation* », à laquelle personne ne s'attendait. Ce sont des réponses populistes aux angoisses. La réponse politique à ces angoisses n'a rien d'évident ; j'espère profondément que nous saurons garantir à la plupart des gens une formation professionnelle continue qui les rendra complémentaires de l'intelligence artificielle.

On nous propose aujourd'hui une utopie techno-marxiste : le revenu universel de base – je précise que le terme de « marxiste » n'est absolument pas péjoratif à mes yeux. C'est une idée absolument suicidaire. Ceux qui le promeuvent, comme M. Benoît Hamon, sont les idiots utiles d'une intelligence artificielle semi-forte. Si nous décidons de mettre de côté tous ceux qui ne sont pas complémentaires de l'intelligence artificielle en nous contentant de leur donner des jeux et du cirque, nous vivrons, dans 50 ans, *Metropolis*, et, dans un siècle, *Matrix* ! Nous devons nous battre jour et nuit, en réformant l'éducation et la formation professionnelle, pour assurer la complémentarité entre les travailleurs, quel que soit leur niveau de qualification, et l'intelligence artificielle faible. À défaut, nous ferions le choix d'un suicide collectif et ce serait de l'irresponsabilité politique et, au-delà, de l'irresponsabilité de toute la société civile. Il y a là un chantier de

plusieurs décennies ; il est regrettable que les politiques, handicapés du mulot qu'ils soient, n'en prennent pas la pleine mesure.

Des contre-pouvoirs pourront-ils émerger ? Comment organiserons-nous la régulation, la gouvernance, la police de l'intelligence artificielle, faible d'abord puis semi-forte ? C'est un débat politique immense. Henri Verdier a beaucoup réfléchi sur ces questions ; je partage en grande partie sa vision. Il faut y travailler dès aujourd'hui. Ce n'est pas gagné ! De tels dispositifs fourniront, demain, des services irremplaçables à nos concitoyens - chacun a pu constater le succès immense et totalement inattendu d'« Écho » et d'« Alexa », lancés par Amazon.

On ne peut qu'être inquiet, à l'image des résultats d'un sondage récent qui montrait que 50 % des Français, et 68 % des jeunes, confieraient volontiers leurs données de santé et leur santé aux GAFA. Il faudra de la pédagogie pour organiser notre sursaut face aux GAFA et aux BATX ! Et pour réaffirmer notre souveraineté ! Il est aujourd'hui plus urgent de réfléchir à ces enjeux de souveraineté, à ces enjeux économiques, scolaires et de formation, que de fantasmer sur une intelligence artificielle forte qui, comme le dit très bien Yann LeCun, ne permet pas d'envisager un scénario à la *Terminator* avant 2035.

Mme Dominique Gillot, rapporteure. - Nous sommes sans voix !

M. Laurent Alexandre. - Je me suis pourtant retenu, aujourd'hui !

Mme Dominique Gillot, rapporteure. - J'étais déjà vaccinée !

Mme Marie-Claire Carrère-Gée. - Juste une précision : Laurent Alexandre estime que nous avons sous-estimé le risque pour l'emploi du développement des nouvelles technologies. Je suis évidemment moins compétente que lui dans le domaine des technologies ; je suis néanmoins frappée par tant de certitudes, dans un monde qui me paraît, à moi, très incertain, s'agissant à la fois du progrès technologique lui-même et de sa vitesse de diffusion. Le progrès technologique n'est pas un système autonome ; sa vitesse de diffusion dépend de maints facteurs : la rentabilité économique, mais aussi les choix sociaux ou éthiques. On sait depuis des années automatiser les caisses dans les supermarchés ; mais l'automatisation ne se fait pas à la minute, et n'est toujours pas accomplie. Je pense nécessaire que nous restions collectivement très humble sur les prédictions. Les citoyens et les pouvoirs publics ont leur mot à dire sur la vitesse de diffusion des progrès technologiques, lesquels sont, par ailleurs et par définition, difficilement prévisibles. J'ai exposé clairement les limites de notre étude : nous avons travaillé sur la base des frontières technologiques actuelles.

Mme Dominique Gillot, rapporteure. - Votre étude n'est nullement mise en cause.

Mme Marie-Claire Carrère-Gée. - J'ai présenté les conclusions d'une étude et non mes certitudes ; à côté de moi se trouve quelqu'un qui professe des certitudes.

M. Laurent Alexandre. – J’ai dit que le futur était imprévisible ! Je n’ai pas de certitudes.

Mme Marie-Claire Carrère-Gée. – Je sais aussi que certains ont intérêt à entretenir des idées fausses sur la vitesse de diffusion des technologies : il est plus facile de lever des fonds en promettant un progrès technologique majeur sous deux ans qu’en exprimant des incertitudes.

Mme Dominique Gillot, rapporteure. – Je me félicite d’entendre s’exprimer des avis contrastés ; c’est par le débat que nous avancerons ! J’aurais aimé que Mme la Secrétaire d’État vous entende, Monsieur Alexandre.

Je vais à présent donner la parole à M. Jean-Christophe Baillie. À la suite de vos études, vous êtes devenu, Monsieur Baillie, chercheur en intelligence artificielle au *Sony Computer Science Lab*, puis vous avez participé à différents projets d’entreprise liés à l’intelligence artificielle et à la robotique comme *GoStyle* ou *Aldebaran*. Vous êtes aujourd’hui à la tête de votre propre entreprise, *Novaquark*.

4. M. Jean-Christophe Baillie, entrepreneur (*Novaquark*)

Laurent Alexandre vient de présenter un grand nombre des points dont je souhaitais traiter. Je vais tenter d’ouvrir des perspectives.

Loin de moi toute certitude sur ces sujets ; ce qui me semble intéressant, c’est de lancer des idées, des questions, des points de vue. Je voudrais notamment débattre de l’« utopie techno-marxiste » dont parlait Laurent Alexandre.

Il faut distinguer entre intelligence artificielle forte et intelligence artificielle faible. J’ai récemment écrit un article au titre provocateur, intitulé « Pourquoi AlphaGo n’est pas de l’intelligence artificielle », dans lequel j’appuie cette distinction. L’intelligence artificielle faible, ou ce que certains appelleraient l’informatique avancée, correspond à ce que nous savons faire aujourd’hui : opportunité économique incroyable au potentiel énorme, elle affectera assurément le marché du travail, mais selon une dynamique assez classique de destruction et de création d’emplois, donc sans bouleverser en profondeur le marché du travail comme pourrait le faire l’intelligence artificielle forte. L’un des défis que nous ne savons pour le moment pas du tout relever, c’est celui de l’interaction sociale pertinente, douée de sens, avec d’autres individus, dans un environnement complexe. Les agents d’entretien que je croise le matin accomplissent un travail d’une complexité ahurissante ; les gestes qu’ils effectuent paraissent évidents, mais les problèmes qu’ils résolvent résistent absolument à la robotique, ce domaine qui se frotte à la complexité du réel et, en effet, ne progresse pas de manière exponentielle.

Appelons cela l'intelligence artificielle forte. L'intelligence artificielle connaîtra sans doute un stade intermédiaire avant que nous soyons capables de créer une conscience artificielle. Je pense néanmoins que nous pouvons arriver jusque-là, mais je suis incapable de donner une date - ceux qui se prêtent à ce jeu pratiquent un exercice périlleux ! Des avancées scientifiques profondes sont nécessaires : il y va de l'élaboration d'une véritable théorie de l'intelligence, rien de moins. Et nous ne savons pas s'il y en a pour dix ou pour cent ans ! Dire le contraire, c'est prétendre qu'il eût été possible de prévoir dès le XIX^e siècle que la théorie de la relativité serait un jour formulée. On ne peut prédire l'émergence d'une idée nouvelle.

Quoi qu'il en soit, le jour où l'on en viendra à une intelligence artificielle forte, on assistera à une modification profonde du marché du travail : les machines que nous fabriquerions seraient capables d'exercer y compris des métiers manuels, grand défi de l'intelligence artificielle.

On peut y voir une catastrophe, la promesse d'une explosion du taux de chômage. Mais ne serait-ce pas là, plutôt, une opportunité pour repenser l'organisation de la société ? Aujourd'hui, notre société est crispée sur la notion de travail, dès l'école. Nous sommes invités, dès l'enfance, à apporter notre pierre à l'édifice commun en trouvant notre place sur le marché du travail. En l'absence de robots, nous n'avons pas le choix ! Quelqu'un doit bien faire le travail, extrêmement important, de l'agent d'entretien. Nous avons besoin d'une société qui valorise le travail et qui, par exemple, à supposer qu'elle mette en place un revenu universel, doit l'établir à un niveau intenable, indécent : si ce niveau est fixé à 2 000 euros, qui acceptera de faire le travail d'entretien dont je parlais chaque matin ?

Pour les besoins de la discussion, on pourrait introduire une distinction entre travail et activité : le travail, c'est ce que vous faites pour subvenir à vos besoins, et que vous arrêtez de faire si vous gagnez au loto - c'est ce que feraient la plupart des agents d'entretien. Par contraste, on peut introduire la notion d'activité : ce que vous aimez faire par passion - si vous gagnez au loto, vous continuez à le faire. Songez, de ce point de vue, à une société dans laquelle la robotique et l'intelligence artificielle forte mettraient fin au travail et autoriseraient la mise en place d'un revenu universel décent, non contradictoire avec la nécessité que les travaux que j'évoquais soient effectués. Dans cette société sans travail, les robots accomplissent ces tâches dont personne, en réalité, ne veut.

Cette société de l'activité ne serait pas une société de l'oisiveté ; simplement, les gens y seraient pleinement libres de décider de l'organisation de leur temps, et y compris, d'ailleurs, de ne rien faire, ou de passer leur vie à la plage en lisant La Pléiade. Une telle activité ne ferait l'objet d'aucune stigmatisation, au contraire de ce qui se passe aujourd'hui, dans une société du travail - une société d'avant l'intelligence artificielle forte -, qui doit inscrire dans sa logique la nécessité de travailler - le chômage, c'est mal ! Une autre logique sociale est imaginable.

Le point clé de cette vision utopique, c'est évidemment l'éducation : il s'agirait de créer des citoyens capables de jouir réellement de ce temps et de cette liberté. Nous n'en sommes pas là ! Si, du jour au lendemain, une grande partie de nos concitoyens étaient mis au « chômage », je ne suis pas certain que les choses se passeraient bien. Il faudra donc créer, *via* l'éducation, un terreau où seront cultivées d'autres valeurs et d'autres visions de la société, ouvrant, à très long terme, une nouvelle ère pour l'humanité.

Pour l'heure, on en reste toujours à l'équation « intelligence artificielle forte égale chômage et fin du travail ». C'est mal, dit-on ! Certes, dans la société d'aujourd'hui ; mais il est possible d'imaginer une société différente.

Mme Dominique Gillot, rapporteure. – Que faites-vous de la dignité sociale, de l'utilité sociale ?

M. Jean-Christophe Baillie. – Il existe certes un besoin communément partagé de se sentir utile ; mais un tel besoin peut très bien se réaliser dans une activité librement choisie, par contraste avec un travail pénible que vous ne pouvez librement abandonner, et qui s'apparente à une privation de liberté. Je ne prône absolument pas l'idée d'une oisiveté généralisée ! J'introduis l'idée d'une activité librement choisie, en accord avec des passions et des intérêts, en un sens très large, nourris par l'éducation : sport, art, culture, accompagnement des personnes âgées par exemple.

M. Laurent Alexandre. – Les robots le feront mieux que nous !

M. Jean-Christophe Baillie. – Je n'en suis pas sûr ! L'idée que le travail serait nécessairement bénéfique au motif qu'il donnerait aux gens une utilité est une erreur de logique.

M. Gilles Dowek. – Ce désir d'utilité sociale est une invention du XIX^e siècle ! C'est une construction sociale. Casanova n'est utile à rien ; il n'en est pas du tout malheureux.

Mme Dominique Gillot, rapporteure. – Je vais à présent donner la parole à M. Jean-Claude Heudin. Enseignant-chercheur, vous avez, dans les années 1990, co-fondé une société spécialisée en intelligence artificielle. Vous avez été expert auprès de l'Union européenne et conseiller scientifique pour la Cité des sciences. Surtout, vous êtes, depuis de nombreuses années, directeur de l'Institut de l'Internet et du multimédia, placé au sein du pôle universitaire Léonard-de-Vinci. Vous avez publié de nombreux ouvrages et vous vous intéressez aux questions politiques et sociétales que pose l'irruption des machines intelligentes.

5. M. Jean-Claude Heudin, directeur de l'Institut de l'Internet et du multimédia

Je vais répéter des choses qui ont déjà été dites ce matin et cet après-midi ; mais il existe un algorithme, en intelligence artificielle, qui s'appelle l'« apprentissage par renforcement ». Je vais donc modestement tenter de renforcer quelques points.

Ma présentation s'articulera autour d'une idée simple : dans un contexte historiquement très anxieux, où il se dit beaucoup de choses sur l'intelligence artificielle, notre priorité devrait principalement aller à l'éducation et à la formation. Le grand public est pour l'essentiel informé par les *blockbusters* américains, qui véhiculent une vision de l'intelligence artificielle complètement déformée, très alarmiste. En tant que praticien de l'intelligence artificielle, je peux vous dire que la réalité, dans les laboratoires, est beaucoup plus laborieuse ! Je vais donc m'efforcer de remettre de la mesure dans ces débats.

L'intelligence artificielle est une petite partie de l'informatique. Ce dont nous avons parlé aujourd'hui à ce titre renvoie en réalité au *deep-learning*, qui correspond à de spectaculaires progrès récents, mais ne représente que la partie émergée de l'iceberg. Que s'est-il réellement passé depuis 2013 ? Nous connaissons un nouvel « été torride » autour de l'intelligence artificielle et, plus spécifiquement, du *deep-learning*. Avant 2013, nous savions construire des applications à partir des réseaux de neurones, mais les résultats n'étaient pas bons – le taux de réussite, s'agissant de la reconnaissance d'une image de chat, par exemple, n'était que de 50 %.

En 2013 se sont conjugués trois facteurs. D'abord, un peu par hasard, une équipe de chercheurs canadiens a réussi à initialiser des réseaux multicouches, ou « réseaux profonds ». Le deuxième facteur est l'accès aux données, qui explique l'avantage des GAFAs, lesquels possèdent de très nombreuses données – des algorithmes, même médiocres, alimentés de beaucoup de données se révèlent bien meilleurs que de très bons algorithmes dont les données seraient plus pauvres. Troisième facteur : la capacité de calcul des cartes graphiques GPU. Résultat : on sait aujourd'hui construire des systèmes dont les taux de réussite sont voisins de 100 %.

Qu'est-ce que cela change ? Auparavant, pour aborder des problèmes complexes et automatiser la réponse à de tels problèmes, il fallait construire des modèles. Désormais, si l'on dispose des données, on peut les faire apprendre à la machine. De nombreux domaines en sont affectés – je pense à la voiture autonome, aux assistants personnels dans le secteur du commerce et des relations clients, à la santé, à la robotique.

Je reviens sur la distinction entre intelligence artificielle faible et forte, qui me semble un peu légère. Je préfère définir six niveaux : l'intelligence sous-humaine pour des tâches spécifiques – la grande majorité des systèmes conçus actuellement, même dans le cadre du *deep-learning*, reste

sous-humaine, y compris la reconnaissance vocale ; une intelligence équivalente à celle de l'humain, pour des tâches spécifiques toujours ; un troisième niveau, supérieur à la plupart des intelligence humaines, à nouveau pour des tâches spécifiques - c'est le stade où nous en sommes avec un logiciel comme AlphaGo ; l'intelligence supérieure à toute intelligence humaine pour des tâches spécifiques, qui nous fait entrer dans l'utopie ; un cinquième niveau supérieur à l'intelligence humaine pour une majorité de tâches - ce n'est pas demain la veille, loin s'en faut ; enfin, l'intelligence artificielle ultime, dystopie ou utopie - pour moi, c'est jamais !

Est-ce un danger pour l'emploi ? Au contraire, rien n'est moins sûr. Nous sommes seulement capables, aujourd'hui, d'établir des corrélations ; en tout état de cause, les pays les plus robotisés ont les taux de chômage les plus bas. Les études se contredisent parfois. Seule certitude : le raisonnement en termes d'emplois détruits et d'emplois créés est binaire, mais l'intelligence artificielle modifie en profondeur un nombre croissant de métiers.

Je précise que la solution n'est de toute façon pas une nouvelle taxe. Ce n'est certainement pas en mettant un boulet aux pieds de nos entreprises et de nos start-up que nous améliorerons notre compétitivité. J'espère ne pas être d'accord avec Laurent Alexandre lorsqu'il dit que nous avons déjà perdu la bataille ! Je suis un indéfectible optimiste : j'ai encore un espoir.

M. Laurent Alexandre. - Je n'ai pas dit que c'était perdu ! J'ai dit que c'était « ric-rac », et qu'il y avait le feu au lac !

M. Jean-Claude Heudin. - Il y a au moins un point de désaccord entre nous : l'intelligence artificielle est un atout français. Certes, Yann LeCun est parti chez Facebook mais la recherche française est de tout premier plan, des entreprises dynamiques existent ; il faut de toutes nos forces les encourager, car les enjeux sont réels.

Il existe néanmoins un déficit d'information du grand public ; les informations qui circulent sont très anxiogènes. Je fais presque une conférence par semaine pour tenter d'évangéliser, de rassurer le public et de démystifier le sujet. Un effort est nécessaire au niveau des formations. Les formations de l'enseignement supérieur ménagent très peu de place à l'intelligence artificielle, ne serait-ce qu'une initiation à ses enjeux. Quelques écoles le font, mais elles sont très peu nombreuses - et je ne parle pas des universités. Et les applications de l'intelligence artificielle dans l'enseignement sont malheureusement quasi inexistantes.

Pour paraphraser Michel Serres, notre société connaît une complexification extrême ; nous allons avoir besoin de toutes les intelligences pour relever les grands défis qui vont se présenter à nous. L'intelligence artificielle doit augmenter notre intelligence, et non la remplacer. Je refuse le vieux démon, qui remonte aux Grecs, de l'annihilation du travail au profit d'une société du loisir. Le travail est très structurant ! J'aime la métaphore de l'intelligence artificielle comme « troisième

hémisphère » de notre cerveau biologique. Elle est indispensable pour la résolution de problèmes complexes qui, aujourd'hui, dépassent notre intelligence biologique. La réflexion éthique, en la matière, est absolument indispensable.

Mme Dominique Gillot, rapporteure. – Vous venez de recentrer notre débat sur la question, déjà abordée ce matin, de l'éducation et de l'information du public, de la construction d'une opinion éclairée.

Je vous pose une question avant de lancer le débat. Quels sont vos liens avec les médias ? Des journalistes spécialisés travaillent-ils sur ce sujet ?

M. Jean-Christophe Heudin. – Je suis un habitué de « *La Tête au carré* », sur France Inter. Certains médias, heureusement, font leur travail ! Mais ces sujets restent anxiogènes pour le grand public. Dans l'enseignement supérieur, et même dans les grandes écoles de commerce, très peu de choses sont organisées sur cette question. Il est pourtant indispensable que les jeunes comprennent l'impact de ces technologies sur leurs futurs emplois. Je le dis souvent : à l'ère des données, du temps réel, de l'analyse prédictive, le *marketing* à la papa est mort.

Mme Dominique Gillot, rapporteure. – Ces derniers mois, nous constatons néanmoins une prise de conscience sur la nécessité d'échanger autour de ces questions : des conférences et des rencontres, comme celle d'aujourd'hui, sont organisées.

6. Débat

M. Jean Ponce. – Je suis enseignant-chercheur. Je ne suis ni économiste, ni politique, ni futuriste – Dieu m'en préserve ! Je voudrais réagir à certaines déclarations péremptoires de M. Alexandre. On ne peut les laisser passer. À entendre qu'il vaut mieux de mauvais algorithmes et de bonnes données, Yann LeCun, que je connais bien, et qui est le pape de ces questions, frémirait ! Ce n'est tout simplement pas vrai.

Je frémis également en entendant dire que la France est un pays du tiers-monde dans ce domaine, ou que la robotique ne décolle pas à cause de son coût. Ce n'est pas vrai – pensez à la voiture autonome ! Je donne quelques exemples concrets d'investissements qui sont réalisés actuellement dans le domaine de la robotique : Dyson, le fabricant d'aspirateurs, vient d'investir cinq millions de livres dans un laboratoire de l'*Imperial College* de Londres ; Foxconn s'apprête à installer 50 000 robots en Chine ; Toyota a investi 50 milliards de dollars au titre de la recherche en robotique dans la *Silicon Valley* ; le gouvernement américain vient de déployer 250 millions de dollars dans un partenariat public-privé sur ce sujet, à Pittsburgh. Le plus grand succès public en matière de robotique est le robot *Roomba*, ce petit aspirateur qui coûte 100 euros. L'histoire des robots chers qui ne décollent

pas, c'est un mythe ! Même chose pour l'intelligence artificielle forte. Il faut faire un petit peu attention à ce qu'on dit.

J'ai apprécié l'intervention de Jean-Claude Heudin ; je reviens simplement sur un détail technique : il est faux de dire qu'en 2013 a été découvert un moyen d'initialiser des réseaux de neurones. On a simplement déployé davantage de données *via* d'importantes capacités de calcul.

M. Jean-Claude Heudin. – Nous pourrions en discuter !

M. Jean Ponce. – Aujourd'hui, nous ne disposons d'aucune explication théorique des raisons pour lesquelles les réseaux de neurones fonctionnent, c'est-à-dire donnent, dans un certain nombre de domaines, d'excellents résultats.

M. Laurent Alexandre. – J'ai été mal compris. Je n'ai pas dit qu'il n'existait pas de potentiel économique et technologique, en France, dans ces domaines ! J'ai dit que nous étions importateurs nets et massifs. C'est incontestable ! La balance de ce que nous importons et de ce que nous exportons, en matière d'intelligence artificielle, est totalement déséquilibrée. *Android* est propriété de Google ; *iOS*, d'Apple. Ce sont deux sociétés californiennes, et 99 % de l'intelligence artificielle consommée en France arrive à travers ces deux plateformes.

M. Jean Ponce. – *Android* est un système d'exploitation ; pour autant que je sache, ça n'a rien à voir avec de l'intelligence artificielle.

M. Laurent Alexandre. – Les applications qu'on utilise tous les jours passent par ces deux plateformes. C'est de l'intelligence artificielle, même si les gens ne l'y rattachent pas spontanément ! Nous sommes importateurs nets d'intelligence artificielle.

Sur la robotique, je n'ai pas été compris. Je n'ai pas dit qu'on ne faisait pas de progrès en robotique. Mais un investissement de 250 millions de dollars, c'est très peu, par comparaison avec les GAFA et l'intelligence artificielle, où les sommes se comptent en dizaines de milliards de dollars. J'ai simplement dit que les progrès, s'agissant de la partie mécanique des robots, étaient moins exponentiels que ceux qui concernent la partie relevant de l'intelligence artificielle. Car, en termes de coûts, dans un robot, il y a plus de mécanique que d'intelligence artificielle et leur taux de croissance en est donc diminué d'autant. Le prix des pièces mécaniques en titane ne diminue pas de 50 % tous les douze mois.

M. Jean Ponce. – Vous plaisantez ! Le fait de parler de pièces mécaniques en titane n'est pas un argument. Je ne mets en doute ni votre enthousiasme, ni votre rhétorique, ni vos connaissances, mais c'est un mauvais argument !

M. Laurent Alexandre. – Il y a des métaux rares dans les robots. Prenons le cas d'Atlas. La robotique, parce qu'elle est en partie mécanique, croît moins vite que ce qui est purement immatériel. En termes de coûts relatifs, c'est évident. La partie mécanique ne suit pas la loi de Moore.

Mme Laurence Devillers. – Votre vision est très restrictive ! Le champ de la robotique est très large. Certains domaines de la robotique se développent plus vite que d'autres.

M. Jean Ponce. – La loi de Moore s'applique uniquement à des aspects mécaniques puisqu'il s'agit des circuits. Elle ne s'applique absolument pas au développement de l'intelligence artificielle, ni, d'ailleurs, à la fabrication électromécanique du robot.

M. Jean-Christophe Heudin. – Gordon Moore est le premier à dire que sa loi ne s'applique pas à l'intelligence artificielle.

D'une manière générale, je suis très réservé dès qu'on parle d'exponentielle. Pour reprendre les termes de Stephen J. Gould, je pense que l'histoire des découvertes scientifiques et technologiques relève plutôt de la théorie des « équilibres ponctués », avec de longues phases de tâtonnement et de soudaines avancées.

M. Jean-Daniel Kant. – Je voudrais revenir sur le thème de l'impact de l'intelligence artificielle sur l'emploi. Il faut être très prudent sur les projections. La question est complexe, et ceux qui assènent des tendances définitives à vingt ou trente ans prennent le risque de se tromper.

Les observations ne donnent aucun résultat clair. Ce sont, disent certains, les emplois peu qualifiés qui seront touchés ; d'autres leur rétorquent que ce sont plutôt les emplois intermédiaires qui sont menacés par l'automatisation, ce que semble confirmer l'exemple californien. En l'absence d'une intelligence artificielle forte, comme l'a dit Jean-Christophe Baillie, il faut des gens pour entretenir les jardins, garder les enfants ou faire le ménage.

Je compte étudier cette question scientifiquement en utilisant les techniques de simulation avec des systèmes multi-agents, dont j'ai parlé ce matin. Seule certitude : ça va transformer l'emploi – de quelle façon ? On ne sait pas – et il faudra reposer la question de la place du travail dans la société. Ce n'est pas rien.

M. Benoît Le Blanc. – Je voudrais apporter trois remarques sur le débat en cours. L'intelligence artificielle donne lieu à de considérables investissements financiers, c'est indéniable. Mais comment savoir s'il ne s'agit pas d'une bulle ? Cette dynamique n'est-elle pas simplement due au fait que certaines personnes ont beaucoup d'argent disponible ? On ne sait pas ce que ça va donner !

Par ailleurs, s'agissant du rapport qu'ils entretiennent avec le travail, des notions de carrière ou d'investissement personnel, les jeunes ont des réactions très inattendues. Ce sont eux qu'il faudrait interroger sur l'évolution du marché du travail.

Enfin, on parle beaucoup, ici, d'exponentielle. La cognition humaine a beaucoup de mal à appréhender certains concepts mathématiques, comme

l'exponentielle, mais aussi, par exemple, les notions liées aux faux positifs et aux faux négatifs. Ces choses-là n'existent pas dans la nature !

M. Olivier Guillaume. – S'agissant des supermarchés sans caisses, les technologies sont en train d'être mises au point. Elles seront utilisables rapidement. Monsieur Alexandre, vous avez dit que l'intelligence artificielle serait meilleure que les radiologues d'ici à 2030, soit dans treize ans ; mais ce sera le cas, d'ores et déjà, dans trois mois, pour le dépistage des cellules cancéreuses, notamment. Nous travaillons sur ce sujet actuellement.

Mme Dominique Gillot, rapporteure. – Ça ne veut pas dire que nous n'aurons plus besoin de radiologues ou de médecins !

M. Laurent Alexandre. – Ce n'est pas ce que j'ai dit.

M. Henri Verdier. – Nous avons beaucoup parlé du travail, je ne pensais pas que nous en parlerions autant. Je le pense depuis longtemps : la révolution numérique sera une révolution organisationnelle et managériale ou ne sera pas. Pendant un siècle, une pulsion de robotisation du travail s'est imposée : obéissance érigée en principe cardinal, instructions très claires et non ambiguës, respect des plans et de la gouvernance, etc. Si le travail consiste à exécuter la pensée d'un comité exécutif, il est robotisable, ou « intelligence-artificialisable ».

Mais si le travail, comme dans la recherche, comme dans les start-up, comme en politique, signifie plutôt l'engagement, la détermination, la créativité, alors il est loin d'être menacé par l'intelligence artificielle. Certaines philosophies du *management* et de l'organisation du travail seront davantage menacées de substitution que celles qui tirent parti du meilleur des humains.

Mme Marie-Claire Carrère-Gée. – Je partage entièrement ce point de vue. Au Conseil d'orientation pour l'emploi, nous travaillons sur l'impact quantitatif et qualitatif du développement de l'intelligence artificielle sur l'emploi, mais aussi sur l'organisation du travail et les conditions de travail – ce sera l'objet du second tome. À quelles conditions la nouvelle vague de progrès technologique est-elle de nature à rendre l'activité ou le travail humains plus riche et moins pénible ?

M. Laurent Alexandre. – L'impact sera très hétérogène en fonction des secteurs. Mais l'avenir est très difficile à prédire en matière d'intelligence artificielle – n'oublions pas que nous avons connu deux ou trois « hivers » de l'intelligence artificielle, pendant lesquels nous avons fait très peu de progrès.

Certaines tâches nous semblent très complexes, comme l'analyse d'un scanner, mais se révèlent relativement faciles à automatiser. S'agissant du diagnostic du cancer du sein, le radiologue est déjà égalé, voire dépassé. L'échéance de 2030, dont je parlais tout à l'heure, c'est la date à laquelle il est imaginable que le radiologue soit dépassé par l'intelligence artificielle pour

tout type d'examen radiologique. Cet horizon est relativement court pour une spécialité qui ne s'imaginait pas du tout menacée il y a dix ans.

Quant à l'intelligence artificielle forte, je n'y crois pas avant plusieurs siècles. Je vous incite à lire un magnifique papier du *New York Times* du 29 juillet 1997, après la victoire de l'ordinateur sur Kasparov aux échecs. Le journaliste interroge un spécialiste de l'intelligence artificielle : « Les échecs, c'est très facile ! Le vrai défi, c'est le jeu de Go ; en la matière, je vous le dis, l'ordinateur ne saura pas jouer avant 100 ou 200 ans ! ». Comme vous le savez, il y a 6 mois, Lee Sedol, le champion du monde de jeu de go, a été écrabouillé par le programme informatique.

Mme Marie-Claire Carrère-Gée. – Des erreurs de prédiction se sont produites dans l'autre sens !

M. Laurent Alexandre. – Certes !

Mme Laurence Devillers. – La comparaison est nulle et non avenue : on compare un humain avec une machine ; mais derrière la machine se trouvent cent ingénieurs au travail !

M. Henri Verdier. – Le couple humain-machine bat toutes les machines et tous les humains ; cela donne des parties bien plus intéressantes !

Mme Dominique Gillot, rapporteure. – Nous sommes d'accord !

M. Stéphane Pelletier. – Ce sujet est extrêmement complexe, voire insaisissable. Je suis le fondateur de *Dreamstarter*, une plateforme *web* qui a pour objectif de faire sortir la santé et l'alimentation des sphères spéculatives.

Je rebondis sur le thème de la santé. Nous discutons d'intelligence artificielle ; l'objectif de ces débats est de dégager une orientation réaliste, dans un monde extrêmement incertain. Des talents existent, qui essaient de convaincre les citoyens, à défaut des pouvoirs publics, du bien-fondé de leurs études ; je pense notamment à celles du professeur Benveniste, qui ont été reprises par le professeur Montagnier, sur la mémoire de l'eau. Au lieu de faire des supputations sur ce qu'il adviendra dans deux cents ans, peut-être pourrions-nous mieux considérer ceux qui proposent des approches pionnières et avant-gardistes.

Mme Dominique Gillot, rapporteure. – L'un n'empêche pas l'autre.

M. Stéphane Pelletier. – Sur l'intelligence artificielle, nous devons réussir à déterminer une orientation réaliste et consensuelle. Je souhaiterais que le citoyen soit au cœur des démarches proposées.

Mme Dominique Gillot, rapporteure. – C'est l'un des enjeux de cette journée !

V. INTERVENTION DE LA MINISTRE : AXELLE LEMAIRE, SECRÉTAIRE D'ÉTAT CHARGÉE DU NUMÉRIQUE ET DE L'INNOVATION

Mme Dominique Gillot, rapporteure. - De nos trois tables rondes, ressort la nécessité de concerner l'ensemble de la population, de mettre en place un fil conducteur dès la petite enfance et tout au long de la vie, pour permettre à nos concitoyens de bien se positionner sur les nouvelles technologies nées de l'intelligence artificielle. Il est en effet indispensable de faire preuve de mobilité professionnelle et de compréhension des grands enjeux éthiques et sociétaux posés par ces technologies. Nous sommes par conséquent très intéressés par votre intervention, Madame la Ministre, d'autant que nous avons appris le lancement demain d'une réflexion importante pour définir une stratégie nationale pour l'intelligence artificielle.

Mme Axelle Lemaire, secrétaire d'État chargée du numérique et de l'innovation. - Sur un sujet aussi complexe et difficile, on voit bien combien les responsables politiques doivent faire confiance aux experts. Merci d'avoir organisé cette journée et cette audition, car le sujet de l'avènement des technologies d'intelligence artificielle dans nos sociétés ne doit pas rester l'apanage des seuls chercheurs et ingénieurs. Il est grand temps que la thématique entre dans la sphère publique car elle se trouve au cœur de notre vie et de notre quotidien.

Le champ d'application de l'intelligence artificielle est de plus en plus vaste, comme j'ai eu l'occasion de le constater lors du *Consumer Electronic Show* (CES) de Las Vegas il y a quelques jours. Un grand nombre d'entreprises présentes là-bas avaient intégré des technologies faisant appel à une certaine forme d'intelligence artificielle, et en particulier au *machine learning*. Les applications sont de plus en plus nombreuses dans le domaine de la santé, de la défense, des relations entre les entreprises et leurs clients. À chaque étape, des commentaires décrivent les risques et les interrogations. L'approche française qui semble prédominer est celle de l'appréhension face au risque, puisque 65 % des Français se disent inquiets du développement de l'intelligence artificielle. Comparativement, 36 % des Britanniques et 22 % des Américains expriment la même crainte. Il existe sans doute des différences culturelles d'approche du risque. Devant ce défi, il n'est pas question que la parole des États soit absente. Le Président Barack Obama a fait publier par la Maison Blanche trois documents sur les conséquences possibles et la nécessité de mieux appréhender le défi de l'intelligence artificielle. Ce faisant, il fait appel à la mythologie de la « nouvelle frontière à conquérir » chère à son pays. Ainsi, l'importance d'investir dans l'intelligence artificielle est comparée au programme Apollo. D'autres États, dont font partie le Japon et la Corée du Sud, commencent seulement à

s'intéresser à ces enjeux. À ma connaissance, l'initiative française est la seule en Europe. L'idée part du constat que la France dispose d'atouts considérables pour s'inscrire en leader dans le domaine de l'intelligence artificielle. L'école mathématique française rencontre l'école informatique, de même que la science des données. De manière non formelle, nous avons dénombré près de quatre mille chercheurs français travaillant dans ces domaines et cent formations, dont d'excellents masters. Cette excellence est reconnue à l'international puisque de grands groupes ont ouvert leurs centres de recherche spécialisés dans notre pays. Des entreprises françaises issues ou non du secteur informatique mènent également des travaux de recherche dans le domaine de l'intelligence artificielle. Je pense notamment à CapGemini, Atos, Sopra Steria, Thalès et Safran, mais nous ne sommes qu'au début de l'histoire. L'enjeu pour le tissu économique et industriel français est de réussir à intégrer ce type de technologies sur des modes de production non exclusivement dédiés au numérique ou aux technologies d'intelligence artificielle.

Prenons l'exemple désormais classique du secteur automobile. La carrosserie avec de l'ingénierie civile intègre de plus en plus souvent du logiciel, ce qui soulève des questions éthiques. Cet exemple illustre la transformation que doivent mener nos industries traditionnelles pour produire des biens offrant une valeur ajoutée sur les marchés internationaux.

Le cadre réglementaire et législatif de notre pays a évolué de manière à accueillir de plus en plus ce potentiel de développement par l'intelligence artificielle. Je pense que le choix fait par le Gouvernement de mettre la donnée au centre des politiques publiques en créant le statut juridique des données d'intérêt général et la mission de service public de la donnée, doit permettre de proposer ce carburant de l'intelligence artificielle qu'est la *data*. Tout cela ne peut naturellement se développer sans s'accompagner du questionnement éthique et sociétal. D'ailleurs, la loi pour une République numérique confiée à la CNIL - dont je salue la présidente Isabelle Falque-Pierrotin ici présente - la mission d'animer le débat public autour des enjeux technologiques, en particulier de l'intelligence artificielle. Le sujet commencera par les algorithmes, mais l'idée est bien d'introduire un débat participatif et collaboratif pour sortir des débats d'experts et aboutir à des choix de société.

J'aimerais ici faire passer le message qu'il existe un formidable potentiel de développement économique et social lié à l'essor de l'intelligence artificielle. Systématiquement cependant, ce développement doit s'accompagner d'un questionnement sur les enjeux éthiques et de modes de gouvernance permettant de définir un cadre réglementaire. Peut-être que ce qui a été fait pour la bioéthique devra être reproduit pour l'intelligence artificielle. Il faut poser les enjeux sans langue de bois et sans jargon. Le but est de prévenir les risques pour se prémunir contre d'éventuels dérapages, et être maîtres de nos choix. Le mot essentiel est en effet la « maîtrise ». Voulons-nous que la France soit uniquement un pays

consommateur au niveau international, ou un pays offreur ? Telle est la question.

Les enjeux éthiques ne sont pas toujours posés sans caricature. Aux États-Unis, le panel des réponses est très large. Les questionnements portent notamment sur les robots tueurs, dont Bill Gates et Elon Musk réclament l'interdiction. Stephen Hawking participe également à ce débat en demandant une limitation de l'usage de l'intelligence artificielle dans le cadre d'une mission internationale se voyant confier des objectifs précis. Nick Bostrom, chercheur à Oxford, a également beaucoup écrit sur le sujet avec une vision sans doute moins alarmiste. Dans tous les cas où la réalité technologique rencontre le potentiel économique, se pose la question de l'éthique et de l'humain. En termes d'emplois, des adaptations seront nécessaires sans pour autant que se produise la destruction catastrophique d'emplois telle qu'elle est parfois annoncée. Il faut que l'ensemble de notre société s'adapte à la nouvelle donne et à la transformation des tâches. Les plus répétitives seront appelées à disparaître pour se concentrer sur la valeur ajoutée, de sorte qu'il faudra faire monter en compétence l'ensemble de la population française.

Confronté à l'ensemble de ces enjeux, il fallait effectivement que le Gouvernement soit au rendez-vous. Le sujet de l'innovation ayant récemment été ajouté à mon portefeuille, j'ai commencé à travailler dès le mois d'octobre à la définition de ce que pourrait être l'intelligence artificielle pour notre pays. Dès demain, aura lieu le lancement des travaux en groupes de travail, aux côtés de Thierry Mandon, mon collègue en charge de l'enseignement supérieur et de la recherche, autour de certains champs prioritaires. Il apparaissait dans un premier temps important de fédérer l'écosystème français, constitué de chercheurs mais également de penseurs sur les questions éthiques et des entreprises. Parmi celles-ci les start-up font sans doute partie des acteurs majeurs de notre pays. Dans un travail de cartographie informel, nous en avons dénombré au minimum cent, dont certaines ont déjà remporté des succès retentissants et des levées de fonds élevées. À l'université, un grand nombre de post-doctorants veulent également être des entrepreneurs. Nous devons par conséquent étudier les passerelles les plus fluides possibles entre la recherche et l'entreprise, dans un enjeu de maîtrise.

Le deuxième axe est celui de l'industrialisation et du transfert des technologies de l'intelligence artificielle vers les secteurs économiques pour en maximiser les retombées sur notre territoire. Le troisième axe de travail aura trait à la définition du cadre réglementaire et institutionnel qui prendrait en considération les enjeux macro-économiques, sociaux, éthiques ainsi que nos priorités en matière de souveraineté et de sécurité nationale.

Tels sont les trois axes qui guideront les travaux des sept groupes, auxquels s'ajouteront d'autres groupes de travail plus spécifiques. Au total, une centaine de personnes se réuniront dès demain. Ces acteurs disposeront d'un délai court pour réfléchir et donner les outils au Gouvernement suivant

pour alimenter les priorités des chercheurs et des entreprises identifiés. Nous sommes au début de l'histoire, mais la France tient à être présente dans l'écriture de cette page très vaste. Nous devons voir l'intelligence artificielle comme un outil au service de la croissance et de la prospérité sociale et sociétale, pour peu qu'il soit bien utilisé.

Mme Dominique Gillot, rapporteure. - La réunion est-elle ouverte à ceux qui n'auraient pas été invités ?

Mme Axelle Lemaire. - Nous avons d'ores et déjà dépassé la capacité d'accueil de la salle mais j'invite tous ceux qui souhaiteraient nous rejoindre à contacter mon conseiller Nathaniel Ackerman. J'aimerais organiser des journées consacrées à l'intelligence artificielle ainsi que des ateliers collaboratifs dans les territoires.

Le lancement de demain nous permettra, à Thierry Mandon et moi-même, d'annoncer la composition du Comité d'orientation. Le groupe de travail « Cartographie » mesurera les forces en présence en France. Les autres groupes de travail seront respectivement consacrés à la recherche et au développement des compétences, à l'industrie et au transfert de technologies vers le secteur économique (robotique, Internet des objets, énergie, application des technologies hybrides, applications médicales...) Un sous-groupe réunira en outre les réflexions consacrées aux véhicules autonomes, finances et commerce. Les groupes en charge des considérations macro-économiques et éthiques, réfléchiront notamment aux thèmes « souveraineté et sécurité nationale », « social et éthique ». Si certains d'entre vous souhaitent se joindre à ces groupes, je les invite à se faire connaître.

Mme Dominique Gillot, rapporteure. - Je pense que votre annonce de cette réunion d'ores et déjà à guichet fermé suscitera encore plus de curiosité. Je donne la parole à la salle.

Débat

Danièle Bourcier, directrice de recherche émérite au CNRS. - Avec ma double casquette de chercheuse et de juriste sur les questions liées à l'intelligence artificielle, je précise que je ne m'occupe pas seulement du droit de l'intelligence artificielle, mais je considère que les juristes eux-mêmes sont concernés par l'évolution de leur métier due au développement de l'intelligence artificielle. J'étendrai d'ailleurs le droit à l'administration. En France, il y a une tradition de l'intelligence artificielle dans l'administration. Depuis le temps que nous évoquons la réforme de l'État, nous nous apercevons que nous devons parler différemment aujourd'hui. Par exemple, une loi résout-elle un problème ? Non car tel n'est pas le but de la loi. Je m'arrêterai là, mais la matière est extrêmement riche. L'intelligence artificielle aura non seulement un impact sur le droit mais également sur d'autres domaines tels que les sciences humaines et sociales. La formation au

niveau universitaire doit être développée sur ce thème dans nombre de facultés de droit.

M. Olivier Guillaume, président d'O² Quant. - Je suis très heureux de vous entendre sur le lancement de cette stratégie et évoquer les sujets que vous avez précisés. Étant président d'une start-up fournissant des briques de technologie en intelligence artificielle pour les grands groupes dont vous avez mentionné le nom, j'ai besoin d'aide pour croître plus rapidement et effectuer le transfert de technologies. Aidez-moi s'il vous plaît.

M. Jean-Daniel Kant, maître de conférences à l'Université Pierre-et-Marie-Curie Paris-VI. - L'intelligence artificielle peut-elle aider le politique à concevoir de meilleures lois ? Peut-être est-ce le cas. Il existe des technologies permettant de simuler l'effet d'une loi sur la société. Le modèle est certes contestable mais il a le mérite d'exister. Cette application d'intelligence artificielle n'est pas souvent citée mais elle concerne le champ politique et le citoyen.

M. Alexei Grinbaum, CEA Saclay, LARSIM. - Quelle est votre vision européenne de l'intelligence artificielle ? Un pays à lui seul ne disposera pas nécessairement des données suffisantes pour pouvoir concurrencer les États-Unis.

M. Pavlos Moraitis, Directeur du Laboratoire d'informatique de l'Université Paris-Descartes. - Je souhaite appuyer la demande de création d'un groupe de travail intelligence artificielle et droit, car je travaille sur les applications de l'intelligence artificielle dans le domaine juridique.

M. Patrick Albert, P-DG SuccessionWeb. - Pourrions-nous transformer le secrétariat d'État en un ministère à part entière ?

Mme Axelle Lemaire. - Je n'avais pas anticipé la question consacrée à « intelligence artificielle et droit » qui est pourtant fondamentale. J'étais ce matin avec le Bâtonnier de l'Ordre des Avocats de Paris, avec lequel nous avons évoqué le degré d'appropriation des avocats quant à l'évolution de leur métier. Grâce aux algorithmes, des prédictions sur les possibilités de gain d'une affaire seront sans doute possibles, tandis que les avocats apporteront une valeur ajoutée essentiellement humaine sur la compréhension des circonstances de la situation et le choix ou non d'aller au contentieux. Le métier d'avocat va ainsi être transformé par l'intelligence artificielle, qui facilite la prise de décision. Dans la loi pour une République numérique, nous avons tenté de faciliter l'accès aux travaux de recherche et aux publications scientifiques. Les chercheurs se sont ainsi vu ouvrir l'accès aux bases de données administratives en les croisant de façon respectueuse avec les données personnelles. À l'heure actuelle, des travaux sont menés par des économistes pour définir ce que serait un revenu universel. Pour cela, il faut connaître véritablement l'état de la pauvreté en France et pouvoir exploiter différentes données.

J'en viens à répondre à la question posée sur les politiques publiques. Je suis la première frappée par l'irrationalité à la base des décisions politiques. Tout le système fonctionne de façon telle que nous n'intégrons pas les données disponibles pour redonner sa légitimité à la décision politique. L'enjeu est celui de la capacité du politique à se reposer sur les travaux des chercheurs pour jauger de l'efficacité des politiques publiques. Je suis très favorable à la constitution d'un groupe de travail réunissant l'intelligence artificielle, le droit et les sciences humaines.

Par ailleurs, le secteur public doit rester maître de la définition des politiques publiques. Ceci est particulièrement vrai dans les domaines de l'éducation et de la santé. Or lorsque des entreprises privées auront acquis une expertise et un quasi-monopole dans l'offre de services en santé et en éducation, l'hôpital et l'école seront alors mis en cause dans leur capacité à fournir ce type de services publics. Selon moi, un enjeu de légalité est en cause, ce qui passe par une nécessaire montée en compétence des administrations pour intégrer des technologies aujourd'hui totalement absentes des modes de prise de décision. J'ai lancé ce chantier avec la Direction générale de la concurrence, de la consommation et de la répression des fraudes (DGCCRF) à Bercy.

L'intelligence artificielle peut clairement aider le politique, au point que j'ai soutenu un projet d'incubateur d'innovation démocratique. Je viens d'apprendre que Facebook va faire de même avec Sciences Po pour un incubateur d'analyse des politiques publiques.

Monsieur Guillaume, je vous invite à prendre contact avec mon Cabinet dans le cadre de la mission French Tech.

Enfin, sur la question du secrétariat d'État ou du ministère, il appartiendra au prochain Gouvernement de définir l'architecture de l'organisation interministérielle sur les sujets d'innovation, qui doivent être considérés comme une priorité politique absolue. Les derniers débats des primaires n'ont sans doute pas encore évoqué ces enjeux.

Mme Dominique Gillot, rapporteure. - Merci beaucoup Madame la Ministre de vos réponses très précises et argumentées. Il faut affirmer le fait qu'on a besoin d'embarquer tous les publics dans l'appropriation des nouvelles technologies, cela ne doit pas se résumer à un partenariat entre les experts et les pouvoirs publics. Cette préoccupation de démocratisation de la nouvelle manière d'appréhender tous les actes de notre vie quotidienne, doit être partagée avec le plus grand nombre de nos concitoyens pour éviter que le clivage se creuse. Il s'agit d'un enjeu de démocratie pour la France, ayant également trait à sa capacité à exercer sa souveraineté dans un monde en pleine mutation.

VI. QUATRIÈME TABLE RONDE, PRÉSIDIÉE PAR M. CLAUDE DE GANAY, RAPPORTEUR : ENJEUX ÉTHIQUES DE L'INTELLIGENCE ARTIFICIELLE

Claude de Ganay, rapporteur. - Je vous propose d'aborder notre quatrième table ronde consacrée aux enjeux éthiques de l'intelligence artificielle. « *Science sans conscience n'est que ruine de l'âme* », comme l'affirmait Rabelais. L'intelligence artificielle représente à cet égard un ensemble technologique par rapport auquel la réflexion éthique est non seulement nécessaire mais urgente.

Je demanderai aux intervenants de tenter de tenir leur temps de parole. Je donne la parole à Gilles Dowek, qui a une double casquette d'informaticien et de philosophe. Vous êtes directeur de recherche à Inria et professeur attaché à l'ENS Paris-Saclay. Surtout, vous êtes un spécialiste des enjeux éthiques de l'intelligence artificielle. C'est pourquoi nous vous écoutons avec beaucoup d'attention.

1. M. Gilles Dowek, directeur de recherche Inria, professeur attaché à l'ENS Paris-Saclay

L'éthique des sciences et des techniques, objet des travaux de la Commission de réflexion sur l'Éthique de la Recherche en sciences et technologies du Numérique d'Allistene (CERNA), n'a cessé de se transformer au cours du temps. En premier lieu, nos valeurs elles-mêmes se sont transformées. Nous essayons de résoudre nos conflits de façon plus pacifique que par le passé, tandis que des questions éthiques différentes sont apparues à chaque époque. Notre époque, que nous pouvons faire débiter aux années 1930-40 est celle de la révolution informatique parfois appelée la troisième révolution industrielle. Les deux premières révolutions industrielles étaient celles de la transformation de l'énergie, alors que la révolution informatique est une révolution des techniques de l'information et de la connaissance, une révolution cognitive. En cela, elle est comparable à l'invention de l'imprimerie. Selon Michel Serres, cette troisième révolution est une révolution cognitive et de notre rapport à la connaissance.

Les deux premières révolutions industrielles ont consisté à remplacer nos muscles par des machines. Les révolutions des techniques de l'information remplacent nos cerveaux par des machines. Il n'est pas surprenant dans ces conditions que les questions éthiques soient différentes. À ce titre, nous pouvons recenser des questions générales et des questions particulières.

La première de ces révolutions globales tient au fait que naguère, le travail d'un certain nombre de personnes consistait exclusivement à traiter de l'information. C'est le cas, par exemple, des caissiers dans les supermarchés ou des conducteurs de camions. Ces personnes se contentent de traiter de l'information en tournant le volant ou en appuyant sur un bouton de caisse. Tel est également le cas d'autres professions, comme les juristes, les médecins ou les enseignants. Dans mon travail d'enseignant, je ne fais qu'acquérir de l'information pour la transmettre. Les tâches effectuées naguère par des êtres humains peuvent aujourd'hui être épaulées par des machines. La libération du genre humain du travail constitue sans doute une bonne nouvelle, car il est satisfaisant de constater que nous travaillerons moins mais avec une productivité plus grande. De ce fait, tous les systèmes fondés sur la valeur du travail (socialisme, capitalisme...) ne nous seront plus d'aucune utilité pour comprendre comment répartir les richesses au XXI^e siècle. Des oppositions naîtront sans doute, mais elles ne seront pas similaires à celles que nous connaissions au XX^e siècle.

La première question éthique qui nous est posée est donc très globale : comment répartir les richesses après la fin relative du travail ?

En deuxième lieu, la petite quantité d'informations que nous sommes capables d'échanger nous avait conduits naguère à inventer des mécanismes d'expression de manière très économe. Par exemple au XIX^e et au XX^e siècle, nous ne pouvions donner la parole quotidiennement à des populations isolées au cœur de la France, de sorte que nous avons inventé un système pour leur permettre de s'exprimer une fois tous les trois ou cinq ans par la voie d'un bulletin déposé au fond d'une urne. Aujourd'hui qu'il est possible à chacun de se faire entendre sans l'intermédiation d'un élu, cette notion même de représentation des citoyens n'a plus de signification. Par conséquent, il nous faut inventer de nouvelles institutions permettant à un individu de s'exprimer seul. Nous avons évoqué la consultation qui a été organisée à l'occasion de la loi sur le numérique, mais il s'agit d'un petit pas. Il sera en effet nécessaire d'aller beaucoup plus loin dans la réforme des institutions. Peut-être qu'un jour, le Sénat n'existera plus et qu'il sera remplacé par une meilleure ou une plus mauvaise solution. Il nous appartient d'y veiller en vertu de notre responsabilité éthique.

La troisième transformation très globale concerne l'école. L'école obligatoire a toujours eu pour mission de répondre aux révolutions industrielles de l'énergie. Désormais, la révolution industrielle de l'information nécessite de revoir l'école de fond en comble pour aider les élèves à vivre à l'ère de l'information. Il faut donc compléter les enseignements de physique et de biologie par des enseignements de l'informatique, des humanités ou de l'éthique.

S'agissant des enjeux plus spécifiques de la révolution actuelle, le traitement des données scolaires ou médicales qui concernent les personnes permet de récolter un grand nombre d'informations sur ces dernières. Or les machines ont la capacité de regrouper ces informations sur une longue

durée. Les ordinateurs et les systèmes d'intelligence artificielle ont une hypermnésie. C'est pourquoi il faut définir de nouvelles normes pour maîtriser l'ensemble de ces données. Par ailleurs, si l'automatisation des traitements médicaux permet aux personnes âgées de gagner en autonomie, cette évolution risque également de les couper de tout lien humain. De ce fait, la mise en circulation d'une machine remplaçant une infirmière nécessite d'inventer de nouvelles formes de liens et de réfléchir au type de tâches que nous souhaitons déléguer aux robots et aux algorithmes.

Une autre nouveauté tient au fait que de longs calculs permettent d'aboutir à des résultats si longs, que nous ne savons pas les expliquer. Par exemple, les ordinateurs calculent tous les jours la température du lendemain sans que nous puissions comprendre comment ils sont parvenus à leur conclusion. En réalité, nous n'avons plus aucun espoir de retrouver notre capacité d'explication, qui était due à notre infirmité pour effectuer de très grands calculs.

Enfin la dernière question éthique et juridique a trait à la création de textes juridiques encadrant les algorithmes et les robots, qui pourraient devenir des personnes morales et des sujets de droit.

Pour conclure, j'observerai que le terme « intelligence artificielle » a été successivement revendiqué par une série de communautés d'informaticiens sans grand-chose en commun, sinon qu'ils demandent à des machines d'effectuer des tâches qui, si elles étaient effectuées par des êtres humains, feraient appel à leur intelligence. Ces différentes communautés et actions de recherche nous ont appris qu'il n'existait pas une, mais des formes d'intelligence artificielle. C'est un mot utilisé pour désigner des choses différentes. De surcroît, la diversité des recherches en la matière ne fait que refléter la diversité des formes d'intelligence que nous attribuons à l'être humain. Nous pouvons bien entendu, continuer d'utiliser le terme « intelligence artificielle » pour définir ces recherches, mais il me semble que nous éviterions nombre de peurs et de fantasmes si nous lui préférons les termes précis qui correspondent à tel ou tel outil : « apprentissage automatique », « reconnaissance des formes », « traitement de la langue » etc. Il apparaît en effet particulièrement important que les différents domaines de l'intelligence artificielle posent les questions éthiques qui leur sont propres, tout à fait différentes les unes des autres le plus souvent.

Claude de Ganay, rapporteur. - Sans attendre, je passe la parole à Laurence Devillers, professeur à l'Université Paris-Sorbonne. Vous êtes spécialisée dans les émotions dans les interactions entre les hommes et les machines. Vous animez actuellement un groupe de réflexion sur l'éthique de l'apprentissage automatique au sein de la Commission de réflexion sur l'Éthique de la Recherche en sciences et technologies du Numérique d'Allistene (CERNA).

2. Mme Laurence Devillers, professeur à l'Université Paris-Sorbonne/LMSI-CNRS

Il y a une très bonne transition à faire après les propos de Gilles Dowek sur la pluralité d'intelligences artificielles. Je vais évoquer les enjeux éthiques de l'apprentissage machine, présent dans de nombreux systèmes d'intelligence artificielle.

Dans le domaine de la robotique, les machines intègrent un grand nombre de systèmes à base d'intelligence artificielle ainsi que d'autres technologies. Au LMSI, laboratoire du CNRS où j'effectue mes recherches, je travaille avec un grand nombre de partenaires sur le projet Romeo 2, consacré à l'interaction avec les robots pour assister les personnes âgées. Pour ce faire, nous menons un grand nombre d'expérimentations sur le terrain et enregistrons des personnes âgées en interaction avec les robots. Nous constatons ainsi l'émergence de multiples sujets éthiques. Pour gérer le dialogue, différentes techniques d'apprentissage machine sont disponibles. Certaines d'entre elles réalisent de très bonnes performances mais sont opaques. Pour une personne en train d'interagir avec ce type de machine, il est absolument impossible de comprendre ce qui se passe. D'ailleurs, je ne pense pas que nous parviendrons un jour à retracer l'ensemble des opérations. Pour autant, il y a une exigence importante d'éduquer davantage la société sur ces sujets. Pour notre part nous, chercheurs dans l'industrie, devons travailler sur la traçabilité et l'explicabilité des systèmes ainsi que sur les évaluations.

L'autre point important, et qui constitue une vraie rupture technologique, tient au fait que ces machines apprennent en fonction de l'interaction avec l'être humain. Il faut donc mettre des garde-fous car ces machines arriveront demain. Le robot Paro, utilisé auprès des personnes âgées dans le cadre de la maladie d'Alzheimer, présente déjà une utilité certaine et crée une forme de lien social. Il est possible de parler à travers une machine, qui peut représenter un médiateur. De plus les personnes qui utilisent les machines, interagiront elles-mêmes avec les patients, ce qui constituera une forme de lien triangulaire.

Je voudrais maintenant assurer une présentation du travail mené au sein de la CERNA en ce moment, qui débouchera sur un rapport en mars 2017 et qui fait suite à un premier rapport sur l'éthique de la recherche robotique en 2016.

Les défis scientifiques de demain supposent de poser les règles morales et éthiques qui assortiront nos implémentations sur les robots. L'un des buts du chercheur en robotique sociale est d'empêcher un déficit de confiance de la part des utilisateurs mais également, de freiner une confiance probable. Les travaux de recherche permettent en effet de constater que les gens projettent sur la machine des intentions anthropomorphiques sur lesquelles il est nécessaire de les détromper. Il est important d'éduquer les

utilisateurs sur leurs projections vis-à-vis de la machine, lorsqu'ils ont tendance à lui attribuer davantage de capacités qu'elle n'en a. Toutefois, le fait que la machine apprenne de nous va susciter un attachement et créer du lien. Je travaille beaucoup sur cet aspect de coévolution entre la machine et l'homme, étant rappelé toutefois que des liens d'attachement peuvent également se créer avec sa voiture ou sa montre. Néanmoins, les laboratoires travaillent aujourd'hui sur les interfaces du dialogue, encore très compliquées, et sur les habitudes des personnes pour pouvoir les mémoriser.

Un grand nombre d'initiatives actuelles portent sur l'éthique, ce dont je me réjouis. Je citerai à titre d'exemple celle de l'IEEE ou institut des ingénieurs électriciens et électroniciens, dont le projet de charte de décembre 2016 met en exergue des leviers importants. Il est nécessaire d'éduquer les chercheurs, les journalistes, les industriels et les politiques sur l'éthique. Celle-ci consiste en réalité à définir des règles, à les appliquer à un robot, mettre en œuvre des outils pour vérifier leur respect par ce robot et prévoir un cadre juridique.

L'apprentissage machine qui se trouve au cœur des systèmes est en réalité très opaque car l'utilisateur ne perçoit aucunement les capacités de la machine. De ce fait, il est extrêmement important de comprendre, à différents niveaux, qu'il faut pouvoir évaluer les performances, et de vérifier les données utilisées. On parle beaucoup d'intelligence artificielle, mais en ce qui me concerne je préfère employer le terme de « bêtise artificielle ». En fonction de leur nature, les données fournies à la machine pourront finalement aboutir à un système très désagréable, discriminant ou raciste. Par conséquent, nous devons tous être conscients des risques liés à la manipulation des outils. Lorsque des industriels recourent au *deep learning* pour opérer une classification de leurs clients, j'attire l'attention sur le fait qu'il s'agisse d'une sémantique de surface. Par conséquent, il faut se méfier des performances de ces systèmes, qui vont certes apprendre mais pas toujours avec du sens. L'intelligence artificielle possède des capacités qui lui appartiennent en propre, mais à l'inverse l'homme est doté de sens et d'une intelligence cognitive que n'aura jamais la machine. C'est pourquoi il ne faut pas opérer de confusion ni comparer constamment l'homme et la machine. Les chercheurs travaillent à élaborer un système utile pour les humains, mais nullement à copier l'humain.

Au sein de la CERNA, qui réunit des juristes, des philosophes et des scientifiques en informatique, nous avons mis en évidence les concepts renouvelés d'apprentissage, de responsabilité, d'explicabilité des systèmes et d'évaluation : nous rendrons donc public un rapport sur ces aspects en mars 2017. Par ailleurs, nous avons élaboré des préconisations et des questionnements classés en six thèmes :

- les données des systèmes d'apprentissage ;
- l'autonomie des systèmes apprenants ;
- l'explicabilité des systèmes d'apprentissage ;

- les décisions des systèmes d'apprentissage ;
- le consentement dans le domaine du numérique ;
- la responsabilité dans les relations humain-machine.

Les mécanismes de contrôle commencent à émerger. Pour élaborer des outils intelligents et autonomes, nous devons nous munir d'outils de vérification et d'évaluation.

La CERNA s'intéresse également à des préconisations générales telles que la formation des enseignants et des étudiants, les plateformes collaboratives de recherche ou la création d'un institut de recherche sur le risque numérique. Sur ce dernier, il importe que la participation soit la plus large possible, et que les grandes institutions de recherche et les industriels en fassent partie.

En outre, j'indique que mon livre à paraître chez Plon le 23 février 2017 sera intitulé « *Des robots et des hommes, mythes, fantasmes et réalités* ».

M. Claude de Ganay, rapporteur. - Je passe la parole à Serge Abiteboul, ancien titulaire de la chaire d'informatique du Collège de France, directeur de recherche Inria, romancier, militant en faveur de l'enseignement de l'informatique et fondateur d'un blog intitulé *Binaire*.

3. M. Serge Abiteboul, directeur de recherche Inria

J'ai choisi de vous parler d'aspects éthiques dans le cadre des données massives. L'intelligence artificielle est la possibilité pour un logiciel de réaliser une activité qui, chez un humain, demanderait de l'intelligence. Je n'évoque par conséquent pas uniquement le *machine learning* mais l'intelligence artificielle au sens le plus général du terme. Ces logiciels peuvent avoir des actions dont l'importance sur notre vie et nos sociétés est considérable. En corollaire, ils doivent se comporter de manière responsable. À titre d'exemple, un moteur de recherches qui biaiserait ses réponses limiterait notre liberté de choix et pourrait entraîner un déséquilibre du commerce mondial. De même, le lancement d'une fausse nouvelle économique pourrait avoir un impact sur le cours de l'action de la société concernée. Dans ces deux cas, le problème vient de ce que l'information est répercutée sur tous les réseaux et les données sont échangées sans être vérifiées.

Partant de ces constats, nous remarquons que de telles activités intelligentes s'appuient souvent sur la masse de données considérables générées par les réseaux sociaux ou nos téléphones. Grâce à l'intelligence artificielle, nous tentons de créer de la valeur. Les recommandations sur les réseaux sociaux dépendent de la masse de données dont ces systèmes disposent. En définitive, l'accumulation de données et leur analyse ont pris

une importance sociétale considérable. Dès lors, il importe que ces données se conduisent de façon responsable.

Pour prendre un autre exemple, la décision sur une demande de prêt représente une tâche relativement répétitive sur laquelle un humain pourrait avoir des biais considérables alors qu'un algorithme se comporterait en théorie de façon plus juste. De telles questions d'équité commencent juste à apparaître. Il est indispensable au préalable que le décideur fixe lui-même ce qui est considéré comme juste avant de faire travailler l'algorithme. Une telle décision ne relève pas du rôle de l'informaticien mais revêt un caractère politique.

Cet aspect de vérification est donc essentiel. De même, la transparence est une notion cruciale. Il convient par exemple qu'un opérateur précise à l'intention du public, qu'il s'interdit de vendre des données confidentielles auxquelles il aurait accès. Or ces exigences de loyauté et de transparence ne vont pas d'elles-mêmes dans les systèmes d'intelligence artificielle. Sur une plateforme marchande ou un réseau social si les recommandations proposées à l'internaute sont trop réduites, son point de vue risque d'être modifié. En définitive, les choix à opérer sont éthiques, et non techniques.

Pour ma part, je suis optimiste car je pense que les algorithmes peuvent nous permettre d'obtenir de bien meilleurs résultats que par le passé. Il faut cesser de regretter constamment le passé. Par exemple, l'affectation des élèves après le bac fonctionnait très mal avant la mise en place de la plateforme APB. L'utilisation d'un algorithme a abouti à une nette amélioration. Mais il manque encore la transparence. Aujourd'hui, l'algorithme est publié et les critères de choix sont publics. Il devient possible d'améliorer davantage le système dans la transparence.

En définitive, les algorithmes peuvent apporter plus d'équité et de transparence que les humains, à la condition toutefois de consentir des efforts de vigilance. De tels efforts émaneront de l'État et de la réglementation, des associations d'internautes et de la technique. Sur ce dernier point, la technique permettra de gérer les systèmes en les analysant de façon plus efficace, par exemple sur la politique de *privacy*. La technique peut également aider à vérifier si un système se comporte de façon convenable sur le *web*.

Pendant cinquante ans nous avons œuvré avec acharnement pour réussir à travailler sur de gros volumes de données et à les analyser : nous y sommes parvenus aujourd'hui. À mon sens, les années à venir seront déterminantes sur les aspects éthiques. Il faudra tenir compte de ces dimensions dès la conception, pour aboutir à des produits qui seront par exemple *Ethic by design* ou *Fair by design*.

M. Claude de Ganay, rapporteur. - Jean Ponce, je vous passe la parole. C'est votre collègue Francis Bach, indisponible, qui nous a proposé de vous inviter. Professeur à l'ENS, passionné de science-fiction, vous vous

êtes spécialisé sur les applications permettant de rapprocher les récits de science-fiction des capacités réelles des robots. Vous êtes ainsi devenu l'un des plus grands spécialistes de la vision artificielle, c'est-à-dire des technologies permettant aux machines de se rapprocher des capacités visuelles de l'être humain.

4. M. Jean Ponce, professeur et directeur du département d'informatique de l'École normale supérieure (ENS)

Contrairement à mes collègues ici présents, je ne suis pas un spécialiste de l'éthique. Personnellement, il me semble que les machines véritablement intelligentes qui en viendraient à se poser des questions éthiques, restent du domaine de la science-fiction. En revanche, j'évoquerai une problématique connexe dans le domaine de l'intelligence artificielle, ayant trait aux attentes et aux risques suscités.

En premier lieu en termes d'attentes, les voitures autonomes ou les robots assistant les personnes fragiles sont des thèmes récurrents de ces dernières années et vont finir par arriver. Les attentes, nées en particulier d'effets d'annonce des grands groupes industriels, suscitent également des craintes et des fantasmes. Cela correspond à un discours marketing qui se comprend, dans la mesure où ces entreprises visent à accroître leurs profits et donc à exagérer les progrès constatés. C'est le scénario de l'intelligence artificielle *Skynet* dans le film *Terminator*. Je ne pense pas que des machines intelligentes prendront de sitôt le pouvoir politique. De même les annonces fréquentes de systèmes obtenant des résultats qualifiés de surhumains, me font en général sourire. Je ne crois pas que les radiologues seront au chômage d'ici à deux ans, surtout que les systèmes d'intelligence artificielle produisent encore des erreurs, telles que des faux positifs et des faux négatifs. Je comprends qu'un robot humanoïde comme Atlas fasse peur, il est costaud. C'est une merveilleuse innovation technologique mais cela reste quasiment un gros jouet télécommandé.

À ces attentes peuvent correspondre des déceptions, des craintes et des risques. Un propriétaire de Tesla est récemment décédé au volant de sa voiture. Pourtant dans ces circonstances il faut davantage blâmer les attentes que le résultat. La voiture autonome n'est pas encore là, et des personnes seront toujours présentes pour pallier les éventuels dysfonctionnements. Tesla ou Volkswagen parlent d'aides à la conduite, pas de voiture autonome. Dans les taxis automatiques Uber à Pittsburgh on a maintenu une personne humaine au cas où.

Les programmes d'intelligence artificielle ne sont pas parfaits. Ils donneront des réponses mais avec un taux d'erreur non négligeable. Pour les voitures autonomes, il y aura des accidents mortels, moins qu'avec la conduite humaine mais le niveau d'acceptabilité est différent.

En outre, il est souvent difficile de prédire la fiabilité des résultats de ces programmes et de quantifier les incertitudes. L'accident de Tesla est-il dû au système de navigation ou de décision ? La réponse n'est pas évidente et pose des problèmes de responsabilité et de régulation.

Pour autant, en dépit de ces réserves, ma foi dans l'intelligence artificielle est intacte, il ne faut pas en douter.

En matière d'éthique et de données, la protection de la vie privée est d'autant plus difficile que les méthodes modernes sont très gloutonnes en matière de données. De ce fait, les grandes entreprises du *web* disposent d'un avantage stratégique par rapport au reste de l'humanité, même si elles ne sont pas autorisées à partager avec les chercheurs ou à publier les données ainsi récoltées. Cela pose par conséquent la question de l'anonymisation des données. Il conviendra de trouver des solutions d'anonymisation de ces données afin de pouvoir les utiliser.

La question des robots tueurs est un sujet important, qui pose la question de la responsabilité des programmeurs et des décideurs militaires et politiques au-dessus.

Par ailleurs, la question des dilemmes éthiques est difficile à résoudre. Les trois lois d'Asimov sont par exemple très compliquées à mettre en œuvre dans les machines.

En conclusion, je ferai le constat qu'il existe une frontière de plus en plus ténue entre la recherche académique et la recherche privée. Il est impératif que les deux communiquent pour répondre aux besoins de la société.

Claude de Ganay, rapporteur. - Serge Tisseron a été l'auteur en 1975 de la première thèse en bande dessinée. Psychiatre et psychanalyste, vous êtes chercheur à l'Université Paris Diderot et spécialiste de Tintin et des robots.

5. M. Serge Tisseron, psychiatre, chercheur associé à l'Université Paris Diderot-Paris VII

Comment l'intelligence artificielle est-elle perçue ? Quels sont les aspects problématiques dans cette perception et comment faire évoluer les choses ? Ces considérations, loin d'être anecdotiques, n'ont été que peu évoquées aujourd'hui.

Nous avons beaucoup entendu que l'intelligence artificielle forte n'était pas pour demain. Or il apparaît effectivement que la fantasmagorie actuelle est celle d'une intelligence artificielle beaucoup plus grande que celle qui existe en réalité. Je m'aperçois qu'il ne suffit pas de détromper le public pour le faire changer d'avis. En réalité, l'intelligence artificielle faible angoisse, précisément, parce qu'elle est faible, car de ce fait, elle est

manipulée constamment par son constructeur, qui remet à jour quotidiennement le système, chasse les *bugs*, évite les problèmes et donc les procès. L'intelligence artificielle faible est par conséquent perçue comme un moyen par lequel l'utilisateur de la machine n'a pas le contrôle sur cette machine.

Un saut fantasmagique prête alors à penser qu'une intelligence artificielle forte et autonome permettra à l'utilisateur de négocier directement avec la machine, même si elle est méchante. Une mythologie autour du libre-arbitre de l'intelligence artificielle s'est donc créée aujourd'hui pour échapper à l'angoisse d'une intelligence artificielle sous contrôle de son fabricant. De ce fait, il m'apparaît important de travailler concrètement sur l'ensemble des repères permettant à l'être humain de garder une différence nette entre la machine et l'humain.

J'évoquerai rapidement deux situations actuelles. Une grande banque (le Crédit mutuel) fait répondre à ses clients par un ordinateur très sophistiqué tel que Watson sans que les appelants soient informés qu'ils ont affaire à une intelligence artificielle. Si ce chemin est poursuivi, les supermarchés seront bientôt équipés de robots-caissiers sans que les consommateurs fassent la différence avec un être humain. Bien entendu cette éventualité est exclue dans l'immédiat, mais il faut anticiper les situations en travaillant d'ores et déjà sur les problèmes concrets. Dans les démocraties, il est nécessaire que le citoyen connaisse toujours l'origine des messages auxquels il est confronté. Dans cet esprit, il est important que toutes les personnes en contact téléphonique avec une machine en soient conscientes. Un message pourrait aviser le client de la façon suivante : « *Vous êtes en contact avec une machine. Si vous préférez parler à un humain, appuyez sur la touche dièse.* »

Le second problème que j'aborderai, vise la situation dans laquelle un individu est en contact avec une machine en le sachant, mais en pensant qu'elle a des capacités équivalentes à celles de l'être humain. Une telle croyance relève de l'anthropomorphisme. Alors qu'il a été si utile au développement de l'être humain, l'anthropomorphisme des machines risque de se retourner aujourd'hui contre l'humain.

L'intelligence artificielle est le concept qui a été retenu, était-ce un bon choix ? Il faudrait se demander pourquoi c'est ce mot qui a été retenu en 1956. Les fondateurs de la discipline n'ont pas vu qu'il allait poser problème ensuite. Il aurait mieux valu parler d'apprentissage automatique ou d'apprentissage logiciel. De même on parle de « machines autonomes » alors qu'il serait préférable d'employer le terme de « machines automatiques », on parle d'« empathie artificielle », alors que mieux vaudrait parler de « simulation d'empathie », on parle, enfin, de « réseaux de neurones » alors qu'on devrait plutôt parler de « réseaux de composants électroniques ». Il faut réviser notre langage, car il risque d'entraîner quiproquos et confusion. Nous assistons de ce fait à une véritable exploitation par les industriels de ce langage anthropomorphe. Les publicités pour les robots au Japon sont très

significatives de ce point de vue. Derrière l'idée d'un robot autonome, se joue la volonté d'un certain nombre de commerciaux de nous faire oublier que les machines sont toujours reliées à leurs fabricants pour nous rendre service mais aussi, le cas échéant, nous manipuler.

L'État a un devoir d'informer sur l'ensemble de ces points. Le citoyen doit toujours savoir s'il est en présence d'une machine ou d'un humain, et être protégé des publicités mensongères. Par ailleurs, il me semblerait intéressant que les groupes de travail évoqués par Mme Lemaire comptent également des représentants des diverses religions car ils sont des personnes ressources.

En conclusion, je me réjouis de la somme des travaux menés sur l'intelligence artificielle. Il est très important de protéger l'homme des dangers physiques et matériels que les machines automatiques pourraient lui faire courir. Il est tout aussi essentiel de prévoir des directives sémantiques protégeant l'homme des dangers qu'il pourrait se faire courir à lui-même en raison d'une mauvaise appréciation de la réalité des machines automatiques. Pour coller à l'actualité, je dirai que la décision du Parlement européen de parler de « personnes électroniques » ou de « personnes robots » pour évoquer certains robots sophistiqués vient encore ajouter à la confusion. Tout cela rend difficile un regard réaliste sur ce qu'est l'intelligence artificielle.

VII. CINQUIÈME TABLE RONDE, PRÉSIDIÉE PAR MME DOMNIQUE GILLOT, RAPPORTEURE : DÉFIS JURIDIQUES INHÉRENTS AUX USAGES DE L'INTELLIGENCE ARTIFICIELLE

Mme Dominique Gillot, rapporteure. - Cette dernière table ronde va évoquer les défis juridiques posés par l'intelligence artificielle. Mady Delvaux a été retenue à Strasbourg et nous a adressé son intervention, que je vais vous lire.

1. Intervention de Mady Delvaux, députée européenne (Luxembourg - groupe S&D), rapporteure du groupe de travail sur la robotique et l'intelligence artificielle

Je tiens tout d'abord à vous présenter mes excuses pour mon absence, malheureusement je ne puis me défaire de mes engagements à Strasbourg.

Je vous remercie pour votre invitation à cette audition publique sur l'intelligence artificielle. Je travaille depuis deux ans au niveau européen sur le sujet de la robotique, thème qui me tient à cœur. Je suis donc ravie qu'un organe national tel que l'Office Parlementaire d'Évaluation des Choix Scientifiques et Technologiques se saisisse de la question

J'espère que mon rapport, voté le 12 janvier dernier au sein de la commission des Affaires juridiques du Parlement européen, suscitera l'intérêt et des débats dans les différents pays de l'Union européenne.

Mon rapport est le résultat de travaux d'un groupe de travail composé de députés de divers horizons politiques et d'experts. C'est une initiative législative, ce qui signifie que le Parlement européen fait des recommandations à la Commission européenne. La Commission pourra décider de donner suite ou non mais devra justifier sa décision. Pour ma part, j'ai bon espoir que notre appel à un cadre juridique en matière de robotique soit entendu : en effet, des drones aux robots industriels, la robotique et l'intelligence artificielle font de plus en plus partie de notre quotidien. La question principale est de savoir comment répondre aux nouveaux défis juridiques et éthiques liés à ces technologies.

Qu'est-ce qu'un robot ?

La première difficulté du rapport est de définir ce qu'est un robot. Le texte concerne les véhicules autonomes, les drones, les robots industriels, les robots de soins ou encore les robots de divertissement. Il ne se penche pas sur les robots pouvant être utilisés comme des armes. On entend par robot une machine physique équipée de capteurs et interconnectée à son environnement dans le but d'échanger et d'analyser des données. Il faut

s'attendre à ce que la prochaine génération de robots soit de plus en plus autonome en matière d'apprentissage.

Quid de l'engouement médiatique pour la création de la personnalité juridique ou d'un statut légal pour les robots ?

L'émergence de robots de plus en plus autonomes nécessite une réflexion autour de nouvelles solutions. Dans ce rapport, nous demandons à la Commission européenne de se pencher sur certaines pistes.

L'une d'entre elles serait de conférer aux robots une « personnalité électronique » limitée, au moins pour les cas où une compensation est nécessaire. L'idée serait d'assurer que, dans le cas où un robot prendrait une décision autonome dont résulterait un dommage pour une victime, le robot puisse être considéré comme responsable du dommage. Cette notion permettrait de garantir l'indemnisation des victimes sans entraver la volonté d'innovation. Il s'agirait en quelque sorte du même principe que celui dont nous disposons actuellement pour les entreprises. Cette proposition n'est néanmoins envisageable que sur le long terme et, encore une fois, n'est qu'une des possibilités en matière de responsabilité.

Quelles sont les propositions en matière de responsabilité civile en cas de dommage ? Sur qui doit reposer celle-ci ? Le fabricant, le propriétaire ou l'utilisateur ?

Nous sommes devant deux possibilités. Le principe de la responsabilité stricte propose que le fabricant soit responsable car il est le mieux placé pour limiter de potentiels dommages. Si nécessaire, il peut se tourner vers ses fournisseurs.

La seconde option serait de mettre en place une évaluation des risques avec des tests au préalable et une forme de compensation à laquelle toutes les parties pourraient contribuer.

Nous proposons également la création d'un régime d'assurance obligatoire, au moins pour les « gros » robots.

Concernant l'aspect social de la robotique, notamment sur la question de l'attachement émotionnel aux robots de soins, il faut rappeler aux gens que le robot n'est pas un être humain et qu'il n'en sera jamais un. S'il peut montrer de l'empathie, il n'en ressent pas. Nous ne voulons pas de robots qui ressembleraient de plus en plus aux humains, comme c'est le cas au Japon par exemple.

Nous avons donc proposé la création d'une charte visant à empêcher les personnes de devenir émotionnellement dépendantes de leurs robots.

À quel point est-il urgent de légiférer en matière de robotique ?

Pour une fois, nous pourrions établir des principes à l'échelle européenne et un cadre légal commun avant que chaque État membre ne mette en œuvre sa propre législation. Une standardisation en la matière

pourrait également profiter à l'industrie : pour rester à la pointe en matière de robotique, l'Europe doit disposer de normes communes.

Sur la question de la responsabilité, les clients doivent être certains de disposer d'une forme d'assurance en cas de dommage. L'enjeu clé est celui de la sécurité, mais aussi de la protection des données : les robots ne fonctionnent pas sans échange de données, ce qui pose la question de l'utilisation de toutes ces informations.

Les robots pourraient créer des emplois dans certains domaines et en détruire dans d'autres, en remplaçant par exemple des personnes peu qualifiées. Comment résoudre ce problème ?

Je pense qu'il s'agit là du plus grand défi pour notre société et nos systèmes éducatifs. Nous ne savons pas à l'heure actuelle comment la situation va évoluer. Je suppose qu'il existera toujours des emplois peu qualifiés. Les robots ne vont pas remplacer les hommes : ils travailleront en coopération avec eux, en les aidant par exemple à transporter des marchandises lourdes.

Nous demandons à la Commission européenne de suivre cette évolution et d'analyser dans quels secteurs l'utilisation des robots détruit des emplois, pour que nous soyons préparés à tous les types de scénarios.

Notre rapport contient également un point controversé qui concerne l'éventuelle instauration d'un revenu universel et une réforme des systèmes de sécurité sociale. Si de nombreuses personnes perdent leur emploi à cause des robots, il faudra leur assurer une vie décente. De la même manière, si nous ne pouvons plus fonder notre système social sur l'emploi, comment le financer ? Nous proposons de le faire *via* la source de ce changement : en taxant le travail des robots.

Il faut que les citoyens soient les grands gagnants de cette nouvelle ère. Sans assurer que nos propositions soient la panacée, nous voulons étudier toutes les possibilités et ouvrir le débat. C'est pourquoi nous invitons les États membres à réfléchir et je vous remercie de répondre à cette invitation.

Mme Dominique Gillot, rapporteure. - J'en termine donc ici avec la lecture de l'intervention de Maddy Delvaux. Je donne la parole à Mme Falque-Pierrotin, présidente de la CNIL, qui en plus de son rôle de régulateur en matière de protection des données, s'est vu confier un rôle de réflexion et d'animation sur les règles éthiques applicables au numérique. La CNIL sera également associée étroitement à la stratégie nationale pour l'intelligence artificielle qui sera lancée demain par le Gouvernement.

2. Mme Isabelle Falque-Pierrotin, présidente de la CNIL

Bonjour à tous, et merci de votre invitation. Le sujet de la protection des données personnelles dans le contexte de l'émergence de nouvelles technologies d'intelligence artificielle et de robotique a clairement été identifié par la communauté des autorités de protection des données à l'occasion de leur conférence mondiale de Marrakech en octobre 2016. Lors de cette conférence, nous avons assisté à la présentation d'Helen, robot japonais au travers duquel émergeaient déjà pour nous un certain nombre de questions complexes en matière de protection des données.

De prime abord, la protection des données personnelles ne semble pas aller de soi avec l'intelligence artificielle car les principes attachés à la protection des données semblent aller à l'encontre du fonctionnement de l'intelligence artificielle.

En premier lieu, plus l'intelligence artificielle collecte un nombre très important de données, plus elle est efficace. Or, aujourd'hui, la collecte des données personnelles doit répondre au principe de minimisation des données collectées.

De plus, le principe de la limitation de la durée de conservation des données personnelles se heurte au besoin de mémoire des technologies d'intelligence artificielle qui s'enrichissent au fur et à mesure.

En troisième lieu, la sécurité des données est relativement délicate et souvent imperceptible pour l'individu dans cet univers de l'intelligence artificielle qui fonctionne de manière invisible pour lui, sans couture, à partir de capteurs qui opèrent des arbitrages.

La quatrième question qui me semble enfin devoir être relevée, a trait aux droits des personnes. L'individu doit être à même de comprendre un traitement et d'exercer par rapport à lui des demandes délibérées. Or, en matière d'intelligence artificielle, il est très difficile d'exercer des droits car il est souvent difficile de savoir auprès de qui les exercer.

Il résulte de tout cela, qu'en première analyse nous sommes en présence d'un objet, qui relève pour partie de la science-fiction, mais qui remet tout de même en cause la conception humaniste européenne plaçant l'individu au centre de la politique de protection des données personnelles.

Devant l'intelligence artificielle en perpétuelle évolution, nous sentons que la protection de la personne ne va pas de soi. Il ne faut pas pour autant renoncer à l'intelligence artificielle. Par conséquent, il convient de changer de registre de réponse et de décaler celle-ci.

La question principale sous-jacente est celle de l'autonomie de la personne. Pour progresser, il convient cette fois de cerner les questions éthiques relatives à celle-ci. À cet égard, la CNIL est à la manœuvre dès lors que la loi pour une République numérique de 2016 lui a confié la mission de conduire une réflexion sur les enjeux éthiques et les questions de société

soulevés par l'évolution des technologies numériques. C'est donc un domaine nouveau pour la CNIL dont l'objectif n'est pas d'émettre elle-même un avis éthique mais de susciter un processus de discussion collectif que feront vivre tous ceux qui souhaitent y prendre part. Nous avons choisi le thème des algorithmes et de l'intelligence artificielle comme premier thème de réflexion.

Il importe en effet que les organismes concernés soit les initiateurs d'un débat destiné à fixer les contours des questions éthiques. À l'automne, nous restituerons les éléments les plus saillants de ce débat éthique pour, le cas échéant, faire des recommandations de politiques publiques. La réponse que nous entendons ainsi préparer à l'intelligence artificielle, pourrait provenir de l'article 1^{er} de la CNIL - souvent oublié - qui dispose : « *L'informatique est au service des individus.* » Ce principe fondamental et très général a ensuite été décliné en dispositions plus processuelles et de méthode.

Comment pouvons-nous construire une coexistence entre des humains, des robots et des algorithmes ? Comment faire de l'« *humanity by design* » sur le modèle de la *privacy by design* ? Nous possédons des éléments permettant de bâtir un socle. À Davos, IBM a récemment proposé une première charte de bonne conduite concernant les robots. Les initiatives allant dans le même sens pourront être consolidées.

Mme Dominique Gillot, rapporteure. - Je passe la parole à Rand Hindi, membre du Conseil national du numérique et pilote du groupe de travail sur l'intelligence artificielle. Votre structure sera elle aussi étroitement associée à la stratégie nationale pour l'intelligence artificielle qui commencera demain. Je vous remercie par avance de votre intervention, que je vous demande assez courte.

3. M. Rand Hindi, membre du Conseil national du numérique, pilote du groupe de travail sur l'intelligence artificielle, président de SNIPS

Je suis également à la tête d'une start-up en intelligence artificielle depuis quatorze ans. En tant que chercheur et entrepreneur, j'ai donc assisté à l'évolution de ce secteur.

Nous avons beaucoup évoqué la régulation de l'intelligence artificielle, mais je pense qu'il est beaucoup plus intéressant de réfléchir à la régulation de l'accès à la donnée nécessaire à la conception et à l'apprentissage des algorithmes. Dans le domaine du nucléaire, nous avons régulé à la fois les usages et l'exploitation des matières premières radioactives. Nous pourrions imaginer procéder de façon similaire pour l'intelligence artificielle en empêchant à la source de collecter un certain type de données potentiellement négatives.

Une autre question clé dans ce domaine, concerne la *privacy*. En la matière, nous entendons qu'il faut faire un choix entre vie privée et intelligence artificielle. Je pense que c'est un faux dilemme car des technologies de chiffrement nouvelles permettraient aujourd'hui d'effectuer des calculs directement sur la donnée chiffrée. Ces aspects sont étudiés notamment dans le cadre des données de santé, beaucoup trop sensibles pour être exploitées par tout le monde. Je pense que dans les cinq prochaines années le travail sur les données de santé chiffrées sera démocratisé.

Par ailleurs, les intelligences artificielles ne font pas le même type d'erreurs que celles des humains. À titre d'exemple, la voiture Tesla a percuté un camion bien visible, alors qu'un humain ne serait jamais entré en collision avec lui. Même si cette erreur de la machine apparaît très grossière, en réalité elle se produit beaucoup moins souvent qu'une erreur humaine. De même dans la finance algorithmique, les erreurs commises sont très peu fréquentes mais ont des conséquences potentiellement beaucoup plus graves que les erreurs humaines.

En définitive, le taux d'erreur des intelligences artificielles est très faible et imprévisible mais les erreurs sont d'une importance bien plus grande que celles d'un humain. Finalement, nous avons vu récemment émerger de nombreuses recherches destinées à modifier le comportement des intelligences artificielles. Par exemple en altérant quelques pixels d'une image, l'intelligence artificielle ne la reconnaîtra plus alors qu'un humain ne verrait aucune différence. Potentiellement, le *hacking* d'intelligence artificielle peut avoir des conséquences très négatives. Une personne mal intentionnée pourrait par exemple introduire des biais pour modifier très légèrement des données-sources et en obtenir ainsi des bénéfices.

Tels sont, je pense, les risques plus subtils de l'intelligence artificielle, mais très difficiles à détecter aujourd'hui.

Mme Dominique Gillot, rapporteure. - Olivier Guilhem, vous avez développé au cours de votre expérience professionnelle une expertise précieuse sur les défis juridiques de l'intelligence artificielle et de la robotique. Merci de nous en faire part.

4. M. Olivier Guilhem, directeur juridique chez SoftBank Robotics (ex Aldebaran)

Je vous remercie de me donner la parole. Plusieurs fois, nous avons évoqué les risques et les craintes liées à l'intelligence artificielle. En ce qui me concerne, je me suis interrogé sur le régime de responsabilité applicable qui pourrait en découler. Il me semble que le régime le plus probable est celui de la responsabilité du fait des produits défectueux.

La doctrine n'est pas unanime en la matière mais une loi de 1998 transpose en droit français une directive européenne de 1985. À cette époque, l'intelligence artificielle n'avait pas la même importance qu'aujourd'hui. Dans une réponse ministérielle de 1998, il est clairement précisé que l'ensemble de la catégorie des biens meubles étaient concernés par cette responsabilité, ce qui inclut les logiciels et donc, l'intelligence artificielle. En d'autres termes, le producteur est responsable du produit qu'il met sur le marché et doit indemniser la victime. Il existe cependant des cas d'exonération lorsque la victime a elle-même concouru au dommage.

Si l'on s'appesantit sur les spécificités de l'intelligence artificielle, on s'aperçoit qu'il s'agit d'un système apprenant qui puise, pour apporter une solution avec une certaine autonomie, dans une base de connaissances. Cette base, tout comme l'apprentissage de l'intelligence artificielle, sont potentiellement le fruit des interactions avec l'utilisateur et donc, avec la victime. Cette situation pourrait donc aboutir à une exonération totale ou partielle du producteur de l'intelligence artificielle.

Lorsqu'un produit mécatronique (drone, tablette...) est fabriqué, le fabricant qui y intègre l'intelligence artificielle en supportera juridiquement la responsabilité. Ceci explique qu'*in fine*, le producteur de l'intelligence artificielle soit aujourd'hui très peu inquiet. Ceci pourrait nécessiter un rééquilibrage.

L'autonomie n'est en réalité que la retranscription des valeurs morales du développeur, tout comme le droit est la retranscription des valeurs d'une société à un instant donné. À ce titre, cette autonomie doit être encadrée.

Si l'on considère l'intelligence artificielle comme un bien, un produit au sens juridique du terme, il s'agit d'un produit actif. Comme tout produit actif, il nécessite de prendre des précautions, en travaillant notamment sur les notions de transparence, de neutralité, de droit de modification et de séparation des IA.

Enfin, les risques systémiques liés à l'intelligence artificielle ne sont pas négligeables et doivent être évoqués. Ces risques renvoient à des problématiques d'indépendance nationale et à des problématiques assurantielles. Si une société gérant de l'intelligence artificielle venait à disparaître, des pans entiers de l'économie disparaîtraient consécutivement. C'est pourquoi un système de continuité de service *a minima* pourrait être mis en place pour prévenir certains risques systémiques uniquement, et ce de manière à ne pas exonérer totalement la responsabilité du développeur de l'intelligence artificielle.

Mme Dominique Gillot, rapporteure. - Maître Bensoussan, vous êtes connu pour vos positions en faveur de la reconnaissance d'une personnalité juridique des robots. Nous avons voulu vous attendre sur ces propositions, même si elles ne sont pas consensuelles.

5. M^e Alain Bensoussan, avocat, président de l'Association du droit des robots

Je vous propose de rêver et de réfléchir aux « droits de l'homme de l'intelligence artificielle » qui marqueraient l'avènement d'une métamorphose. La France a exporté ses droits fondamentaux, puis ses droits numériques avec la loi informatique et libertés, elle peut aujourd'hui créer des droits de l'intelligence artificielle qui auraient vocation à se diffuser dans le monde. Les droits de la personne-robot ont de l'importance car les robots vivent avec les hommes.

Je vois quatre défis juridiques essentiels à relever.

Tout d'abord, les algorithmes sont partout et les robots le seront de plus en plus, de sorte qu'il faut nécessairement pouvoir identifier les robots logiciels ou les robots physiques.

Il faut ensuite également repenser la responsabilité en biface. Le robot est toujours responsable vis-à-vis de la victime. Pour le reste, il existe une présomption de responsabilité à la condition d'une certification préalable et sous réserve de fraude.

Bien évidemment, un régime d'assurance devra être prévu lorsque les robots seront parmi nous. En effet, ils entreront bientôt de plus en plus dans les écoles et les entreprises.

Enfin la personne-robot possédant une responsabilité et un régime d'assurance propres, elle doit de ce fait recevoir une personnalité juridique singulière, dans laquelle il faudra par conséquent intégrer le principe de dignité des robots.

Bienvenue à la personne-robot reconnue en tant que telle, et qui devra vivre avec les humains. Dans ce cadre, quatre lignes directrices doivent être conduites pour une cohabitation conforme à nos valeurs humaines entre les hommes et les robots.

La première d'entre elles réside dans la nécessité de permettre d'établir un contrat. Il existe aujourd'hui des contrats entre robots mais également des contrats entre robots et humains. Il me semble temps de penser à un contrat normé où la conception, la conformité, la certification, la garantie et la propriété seraient singulières car les droits actuels sont inapplicables à l'intelligence artificielle telle qu'elle a été décrite.

Le deuxième élément de la mixité est celui de la traçabilité des robots. En traçant les robots, on trace aussi les hommes qui sont à l'origine de leur conception. Il faut donc trouver un point d'équilibre.

Le troisième élément est celui du « principe du bon samaritain », que l'on retrouve dans le droit américain et qui consisterait à revisiter en droit civil quelques règles issues des lois d'Asimov.

Enfin le dernier élément de la mixité entre l'homme et le robot réside dans la possibilité de donner la décision ultime à l'être humain. Ce serait donc un bouton rouge.

6. Débat

M. Sébastien Candel, président de l'Académie des Sciences. - J'ai été très intéressé par le débat de ce jour. J'ai beaucoup apprécié l'intervention de Gilles Dowek, mais il ne faut pas oublier que nous n'avons pas résolu tous les problèmes liés à l'énergie. Les aspects relatifs à la production d'électricité, notamment, sont encore d'une actualité brûlante. La deuxième révolution industrielle continuera d'être un problème pour l'humanité à l'avenir et ne sera pas réglée par des solutions purement informatiques. De plus, nous ne sommes pas parvenus à l'ère de la fin du travail, qui ne disparaîtra pas par le biais de la troisième révolution industrielle.

M. Gilles Dowek. - Je plaide non coupable. Sur les questions de production d'électricité, j'irai dans votre sens. La production d'électricité à partir d'énergies nouvelles sera également cruciale. Je n'ai jamais prétendu que tous les problèmes étaient résolus.

M. Alexei Grinbaum, CEA-Saclay, LARSIM. - Un aspect n'a pas été assez développé, me semble-t-il. L'homme qui interagit avec une intelligence artificielle, change lui aussi. L'homme des années 1900 et celui des années 2000 est bien différent. De plus quand nous évoquons les normes, les valeurs et les droits, le sens des grands concepts qui désignent les droits de l'homme change également. La dignité en 2050 n'aura peut-être pas le même sens qu'aujourd'hui. Nous devons tenir compte de cet aspect important.

Mme Dominique Gillot, rapporteure. - Nous sommes en effet dans un mouvement perpétuel.

Mme Gaëlle Marraud des Grottes, journaliste *Revue Lamy de droit civil*. - Nous avons évoqué très rapidement le concept de transhumanisme et la logique assurantielle.

M. Olivier Guilhem, directeur juridique chez SoftBank Robotics (ex Aldebaran). - Effectivement, nous ne parlons pas autant de transhumanisme en France qu'aux États-Unis. Il convient d'être prudent sur ce mouvement de pensée. Le transhumanisme est un outil marketing. Un outil marketing très puissant car il cherche à donner du sens et une vision.

Mme Laurence Devillers, professeur à l'Université Paris-Sorbonne/LMSI-CNRS. - La façon dont vous en parlez apporte déjà

beaucoup de crédit à ce mouvement philosophique non reconnu qu'est le transhumanisme.

M. Olivier Guilhem, directeur juridique chez SoftBank Robotics (ex Aldebaran). - Je suis entièrement d'accord.

Mme Dominique Gillot, rapporteure. - C'est pourquoi nous n'avons pas inscrit le transhumanisme pour le moment à nos travaux. Nous serons peut-être contraints d'y réfléchir par la suite.

M. Olivier Guilhem, directeur juridique chez SoftBank Robotics (ex Aldebaran). - Nous devons en discuter cependant afin de pouvoir le contester.

Mme Dominique Gillot, rapporteure. - Nous avons délibérément choisi de ne pas l'évoquer pour nous concentrer davantage sur la compréhension de l'intelligence artificielle.

M. Gérard Sabah, directeur de recherche honoraire au CNRS. - Nous devons faire attention à ne pas trop promettre. Nous sommes à l'aube d'une évolution encore embryonnaire, sur laquelle nous devons être prudents pour ne pas susciter de déception. Il faut faire attention à ce que les promesses actuelles n'aboutissent pas à des phénomènes de rejet comme pour la reconnaissance vocale. Nous devons aujourd'hui apaiser le *buzz* pour revenir à un débat beaucoup plus serein.

Mme Dominique Gillot, rapporteure. - Merci pour ces recommandations qui vont dans le sens de notre travail, essentiellement à des fins pédagogiques autour de l'intelligence artificielle.

VIII. CONCLUSION

Mme Dominique Gillot, rapporteure. - À titre personnel, je tiens à remercier l'ensemble des participants d'avoir apporté leur contribution sans partir dans de grandes controverses. Nous saluons également la présence d'Axelle Lemaire et la précision de sa pensée. Les engagements qu'elle a pris pour le Gouvernement aujourd'hui serviront demain la Nation.

M. Claude de Ganay, rapporteur. - Je salue la qualité des intervenants. Nous espérons que l'OPECST contribuera à l'effort pédagogique autour de l'intelligence artificielle grâce à l'ensemble des auditions qu'il mène actuellement.

ANNEXE 8 : COMPTES RENDUS DES AUDITIONS BILATÉRALES CONDUITES PAR LES RAPPORTEURS

I. AUDITIONS DU 25 OCTOBRE 2016

1. M. Stéphane Mallat, professeur à l'École normale supérieure (ENS), chercheur en mathématiques appliquées

Une véritable révolution est à l'œuvre sous les effets produits par le développement du *big data* et de l'intelligence artificielle, dont les conséquences sont encore très mal perçues dans la société civile. À travers l'analyse de données, il est possible de faire des choses qui étaient, pour la plupart des scientifiques, inimaginables il y a cinq ou dix ans. **Cette révolution résulte de la conjonction du regroupement de très nombreuses données dans des bases de données et des capacités de calcul qui ont connu une évolution algorithmique très importante.**

Cette révolution a des conséquences énormes pour l'industrie française. Il y a eu, dans un premier temps, une forme de retard en France, l'industrie numérique s'étant essentiellement concentrée aux États-Unis d'Amérique, qui a de fait attiré de nombreux talents français. Il est donc nécessaire pour la France de rattraper son retard à l'amorçage, qui est classique, mais surtout de récupérer des talents au niveau industriel.

Au-delà de l'industrie, **d'autres domaines sont concernés de manière cruciale par cette révolution.** La grande difficulté pour la médecine est que les données ne sont pas disponibles. Cela soulève un **paradoxe étrange : d'une part, toutes les données personnelles mises sur Internet sont librement à disposition pour des utilisations marketing et commerciales ; d'autre part, au niveau de la médecine, il est extrêmement difficile d'obtenir les données alors qu'il y aurait un impact sociétal extraordinairement bénéfique et important à les exploiter.** Ce paradoxe s'explique, au niveau de la médecine, par des problèmes fondamentaux de confidentialité et de sécurité ; cependant, tout ce qui est commercial et financier est laissé libre et évolue très vite, alors que tout ce qui est médical n'avance pas, alors même que ce serait le domaine le plus bénéfique pour la société.

Il est également crucial de prendre en considération l'enjeu de régulation et de recherche d'équilibre dans l'utilisation des données. Les enjeux de protection de la vie privée sont extrêmement importants, et ils vont le devenir de plus en plus. Il est néanmoins nécessaire d'équilibrer le rapport existant car il ne faut pas empêcher le médical de se développer, le principe de précaution freinant grandement le domaine médical alors que tous les autres domaines évoluent très vite.

Les individus restent souvent sur la perception de l'intelligence artificielle des années 1980 qui était fondée sur une idée de règles et de logiques. **L'intelligence telle qu'elle est construite actuellement relève davantage de l'ordre de l'analogique, qui consiste à des mises en relation entre des éléments analysés.** Les problèmes complexes devant lesquels il est possible de se retrouver, que ce soient des problèmes de reconnaissance d'image, de son, de décisions politiques, sont des problèmes qui résultent de l'interaction d'énormément de variables. Ce ne sont pas des règles qui régissent ces variables ; c'est une forme d'agrégation progressive de ces variables qui amène vers la solution la plus probable, à l'image de signaux émis par des milliers d'indicateurs faibles dont l'agrégation permet d'obtenir une information forte. **Il est de fait impossible de savoir pourquoi un système d'information a pris cette réponse ou décision plutôt qu'une autre.** Les réponses des intelligences artificielles sont le résultat de l'agrégation d'énormément d'indicateurs ; c'est bien plus proche de l'analogie que de la logique et des règles.

L'intelligence artificielle est une opportunité pour la médecine, pour l'énergie, pour l'évaluation des politiques publiques, pour le développement d'innovations, etc. Cependant, **la protection de la vie privée devra être défendue de manière intelligente, afin de ne pas bloquer les effets positifs apportés par l'intelligence artificielle.** Néanmoins, la privauté et la liberté individuelle doivent être préservées. Par exemple, il est nécessaire d'éviter que des systèmes de prédiction totalement ouverts mettent les individus dans des cases et les empêchent d'en sortir, tout en encourageant l'ouverture des données à la médecine afin que cette dernière puisse progresser et se développer. Bien que l'introduction de nouveaux types de systèmes à intelligence artificielle implique nécessairement une prise de faible risque, les bénéfices qui en seront retirés au niveau médical sont considérables. Un équilibre juste et bénéfique doit donc être trouvé.

Du côté industriel, une difficulté tend à émerger au niveau national et européen étant donné le fait qu'il n'existe aucun acteur majeur qui stocke Internet. **Les individus et acteurs de l'économie sont conscients de l'ampleur de la révolution numérique, et l'enjeu sera d'avoir des champions industriels.** L'absence d'acteurs français et européens provoque une fuite des cerveaux et des chercheurs vers les États-Unis d'Amérique. La recherche est d'abord valorisée par le tissu de petites start-up dynamiques qui commencent à émerger en France ; celles-ci se font cependant rapidement racheter par des acteurs nord-américains lorsqu'elles émergent. Il est donc essentiel qu'une politique industrielle aide à l'émergence de géants industriels dans ce domaine afin qu'ils puissent absorber la technologie. Les grandes entreprises du numérique, telles que Huawei ou Facebook, s'installant à Paris, parviennent à capter tous les talents français en particulier car il n'y a pas de plus grosse concurrence au niveau national.

2. M. Patrick Albert, entrepreneur (créateur de ILOG), chercheur et pionnier dans le domaine de l'intelligence artificielle

La France dispose d'un écosystème très favorable à l'émergence de jeunes pousses du numérique, notamment grâce au crédit d'impôt recherche, les pôles de compétitivité, les formations et écoles de l'enseignement supérieur, les sociétés d'accélération du transfert de technologies, les incubateurs d'entreprises, etc. Depuis quelques années, il s'est opéré un changement de discours, de communication et d'appréciation sur l'innovation, ainsi que des outils mis à disposition des entrepreneurs, pour favoriser le développement de l'économie numérique en France.

L'accélération du développement de l'intelligence artificielle s'est faite en trois vagues. La **première vague** correspond à **l'apparition des systèmes experts**, qui sont des logiciels de production des raisonnements et réponses à partir de règles et de faits. Cependant, l'informatique n'était pas prête pour ces systèmes, car ces derniers étaient très complexes.

Au cours des années 2000, **un renouveau s'est opéré avec la simplification des systèmes experts, qui sont devenus accessibles et donc commercialisables et utilisables.** Cette vague d'automatisation des décisions est arrivée à maturité à la fin des années 2000, et a donné lieu à de nombreuses acquisitions. Ce sont principalement des sociétés de *B2B*¹, telles qu'IBM ou Oracle, qui ont acheté des sociétés produisant des systèmes automatisés disposant d'un niveau d'intelligence artificielle « symbolique », c'est-à-dire assez simple, ainsi que des logiciels d'optimisation, aux algorithmes plus complexes, qui permettent de prendre des décisions tirant le meilleur parti des ressources à disposition.

Les industries qui se sont « mondialisées » précocement ont eu une pression de compétition féroce, et ont donc dû baisser leurs coûts. Ces industries, notamment dans les transports, ont de fait eu recours à ce type de logiciels afin de faire face à la concurrence en utilisant des techniques de lean management².

La **troisième vague** a été **l'avènement de l'apprentissage automatique** (*machine learning*). Les résultats des recherches en apprentissage automatique, après leur publication, ont été reçus très positivement par le grand public. Le cas du développement de la voiture autonome s'inscrit dans cette perspective : le grand public comprend son fonctionnement, son intérêt et ses enjeux notamment en matière de sécurité, mais également les dilemmes éthiques que l'utilisation de véhicules autonomes pose.

¹ En anglais : business to business. Ce terme désigne l'ensemble des activités commerciales nouées entre deux entreprises.

² Méthode de management qui vise l'amélioration des performances de l'entreprise recherchant les conditions idéales de fonctionnement de manière à ajouter de la valeur avec le moins de gaspillage possible.

La **quatrième vague** pourrait être celle des **agents intelligents, en particulier grâce à l'Internet des objets**. Cette quatrième vague pourrait être **favorisée par le développement de l'intelligence en essaim** (*swarm intelligence*), dont les méthodes se fondent sur les modèles de colonie de fourmis ou d'essaims d'abeilles grâce auxquels des agents peu doués individuellement parviennent collectivement et en s'organisant à des résultats extraordinaires. **Cette quatrième vague incarne un besoin d'automatisation des actions sur Internet**, à l'image de l'apparition de nombreux agents conversationnels (*chatbot*) sur les sites Internet marchands, intervenant lors de l'acte de vente ainsi qu'en support.

L'enjeu, pour les parlementaires, est la réglementation de l'intelligence artificielle, en prenant en compte les problèmes sociaux et sociétaux qui risquent d'être critiques. La maîtrise de l'intelligence artificielle a la même envergure stratégique pour la souveraineté nationale que la maîtrise de l'atome l'était après-guerre. Cependant, l'intervention du politique est également essentielle pour que les bouleversements économiques et technologiques déclenchés par l'intelligence artificielle n'aboutissent pas à des crises sociales fortes, comme ce fut le cas au début de la première révolution industrielle. L'intelligence artificielle provoque une restructuration majeure de la création de valeur, profitant en premier lieu à ceux qui sauront maîtriser l'intelligence artificielle.

II. AUDITIONS DU 8 NOVEMBRE 2016

1. M. Marc Mézard, directeur de l'École normale supérieure (ENS)

Une des grandes évolutions à prévoir sera l'impact des sciences informatiques, de l'intelligence artificielle et de la science des données sur le vaste champ des sciences humaines et sociales, notamment sur les aspects éthiques. De nombreux autres domaines, comme la médecine, seront également concernés par étape par l'intelligence artificielle.

L'intelligence artificielle est un domaine dans lequel la frontière entre la recherche fondamentale et la recherche appliquée dans le monde industriel est très poreuse. Les problèmes posés par les industriels sont des problèmes de recherche fondamentale. Il existe un besoin énorme de « bras » à fournir pour le tissu scientifique et le tissu économique, les établissements d'enseignement supérieur devant relever le défi de fournir ces « bras » à la fois du côté de la recherche et du côté de la formation.

Le problème de l'emploi face aux enjeux de l'intelligence artificielle est à la fois économique et européen. Il est impératif que s'enclenche une dynamique universitaire qui permette d'attirer et de stabiliser des jeunes talents pour alimenter les entreprises européennes du numérique. L'effort fourni actuellement doit s'intensifier concernant

l'intelligence artificielle. **Le développement de l'intelligence artificielle n'est pas une incrémentation ; c'est un véritable changement de paradigme d'un point de vue scientifique et industriel.** En cela, la question du développement des méthodes des algorithmes est fondamentale. Sur ce plan méthodologique, la France est très bien placée. Cependant, de nombreux autres enjeux cruciaux doivent être pris en compte dans la recherche, notamment concernant les données et l'éthique.

Bien que les champs disciplinaires puissent éventuellement être redécoupés compte tenu des changements liés à l'intelligence artificielle et aux mégadonnées (*big data*), il demeure que l'émergence de nouveaux champs disciplinaires est incrémentale et nécessite le développement de nouveaux codes culturels. Cela implique de créer des instituts ou des centres où cohabitent des étudiants et chercheurs de différents champs disciplinaires, permettant une émulation culturelle favorisant l'interdisciplinarité et l'émergence de nouveaux champs disciplinaires.

Cependant, l'émergence d'un nouveau champ disciplinaire ne peut être décrétée ; elle est conditionnée par le temps nécessaire pour la formation. Il est important que les scientifiques restent attachés à leur discipline initiale qui possède ses questions d'intérêts et d'évaluation. Si les scientifiques sont détachés de leur communauté, l'interdisciplinarité et la création de nouveaux champs disciplinaires ne peuvent pas fonctionner. En revanche, l'interdisciplinarité et l'apparition de nouveaux champs académiques pourraient être favorisés par la mise en avant de l'importance du sujet de l'intelligence artificielle et l'élaboration des projets nécessitant des compétences interdisciplinaires, au travers d'objectifs scientifiques et politiques.

Dans cette perspective, **la participation de la puissance publique est essentielle.** Encore une fois, l'apparition de nouveaux champs ne se décrètera pas mais s'effectuera au travers d'initiatives, de projets et de programmes scientifiques proposés notamment par les pouvoirs publics qui devront accompagner ces évolutions majeures.

L'abaissement de la barrière entre recherche fondamentale et recherche appliquée en entreprise est très perceptible dans le domaine des jeunes entreprises émergentes du numérique. La barrière à franchir pour créer une entreprise dans le domaine de l'intelligence artificielle et, de manière plus générale, dans le domaine des sciences et technologies de l'information et de la communication, est beaucoup moins élevée que celles existant dans des processus industriels traditionnels. En cela, il n'est pas rare que des anciens élèves de l'ENS, qui n'est pourtant pas spécifiquement tournée vers le monde de l'entreprise, intègrent des incubateurs et fondent des entreprises tournées vers les sciences et les technologies numériques.

2. M. Raja Chatila, directeur de recherche au Centre national de la recherche scientifique (CNRS), directeur de l'Institut des systèmes intelligents et de robotique (ISIR)

Le fonctionnement de la robotique présente certaines différences par rapport à celui de l'intelligence artificielle. Dans la perspective de l'étude de la robotique et des systèmes intelligents, **la machine est considérée comme matérielle, et non un logiciel effectuant des opérations ou un algorithme fonctionnant à partir de données.** L'utilisation du qualificatif « matérielle » signifie que **la machine physique interagit directement et de manière automatique avec son environnement.** La différence principale existant entre la robotique et l'intelligence artificielle est issue de la complexité de l'interaction avec la réalité qui implique que les problèmes traités ne sont pas totalement abstraits dans un monde virtuel ; la réalité est porteuse de complexité, d'incertitude, de problématiques liées à la perception d'un environnement complexe qui évolue indépendamment du système et de difficultés d'interactions physiques et matérielles.

De fait, la problématique de la robotique est plus vaste que celle de l'intelligence artificielle, puisque la recherche en robotique doit tenir compte de la perception, de l'action et du mouvement. Les machines sont munies de capteurs afin de percevoir l'environnement, et sont également munies d'actionneurs pour se déplacer et agir sur cet environnement. Ces capteurs, qui extraient des données, doivent permettre de transformer ces données en connaissances pour que la machine puisse décider et agir. L'acte de décision et la capacité d'apprentissage sont permis par des logiciels, impliquant que les ordinateurs sont physiquement des parties prenantes de la machine. **Une machine peut donc disposer de capacités de perception, de mouvement, de prises de décision pour décider de ses actions, de communication et d'interaction pour évoluer dans un environnement humain, et d'apprentissage qui lui permettent d'appréhender la complexité du monde.**

Le robot est le paradigme de l'intelligence artificielle « incorporée » - je préfère ce mot à « incorporée », cela correspond plus à l'anglais *embodied*. Le robot a besoin de la prise de décision et des capacités d'anticipation et d'action d'une intelligence artificielle pour construire ses raisonnements sur les moyens et les contraintes qui permettent ses mouvements. Du fait de la complexité du monde, le robot est construit de manière simple et son système peut se complexifier ensuite au fil de son apprentissage.

La recherche en robotique a débuté de deux manières différentes. D'un côté, la recherche en robotique a été motivée par l'industrie, et plus particulièrement l'industrie automobile qui avait besoin de placer sur ses chaînes de production des robots destinés à effectuer des mouvements répétitifs et rapides. De l'autre côté, les premières réflexions menées sur la notion d'intelligence artificielle et la question posée par Alan Turing sur la capacité des machines à penser ont entraîné le lancement de programmes de

recherche en intelligence artificielle et, à partir des années 1960, en « robotique intelligente ». Ce n'est que dans les années 1980 que la recherche en robotique et la recherche en intelligence artificielle se sont distinguées, la robotique s'intéressant davantage à l'interaction de la machine avec son environnement et aux représentations probabilistes. Elles se rapprochent à nouveau dans la période récente.

En matière de sécurité, les robots comprennent toujours un bouton rouge, c'est bien mais c'est déjà trop tard, il faut tout prévoir pour ne pas en arriver là. L'état du système doit être constamment observé et il faut pouvoir détecter toute déviance avant l'arrivée des problèmes.

Sur le projet transhumaniste, il faut voir que derrière ces discours, nous avons des vues de l'esprit qui n'ont rien d'opérationnelles, elles sont en réalité des idéologies, qu'on cherche à imposer pour gommer les différences entre l'humain et le non-humain. En tant que roboticien, je ne peux que m'opposer à cela.

III. AUDITIONS DU 9 NOVEMBRE 2016

1. M^e Alain Bensoussan, avocat, président de l'association pour les droits des robots et M^e Marie Soulez, avocate spécialisée sur les TIC dans son cabinet

Un robot peut être défini comme une machine intelligente capable de prendre des décisions de manière autonome, maîtrisant son environnement et pouvant le modifier. Un robot est un composé d'algorithmes doté de capteurs et d'actionneurs. De fait, il faut distinguer les robots logiciels des robots physiques : les robots physiques sont des robots logiciels qui peuvent modifier leur environnement grâce à leurs capteurs et leurs actionneurs. La majorité des robots actuellement en circulation sont des robots logiciels ; cependant, en 2020, ils auront tous migré dans des coques dotées de capteurs, faisant d'eux des robots physiques.

Il existe un besoin juridique fort concernant le développement de la robotique à usage non militaire au niveau mondial. Certaines voies permettant de répondre à ce besoin furent identifiées dans le projet de rapport de la commission des affaires juridiques du Parlement européen contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique rendu le 31 mai 2016. J'identifie pour ma part les points suivants, qui représentent un caractère prioritaire pour un développement juridiquement serein de l'intelligence artificielle et de la robotique.

- **Identification des robots.** Les robots doivent être facilement identifiables, « immatriculés », par un nom, un prénom, un numéro IP, un numéro de contact, etc. Ce besoin impératif d'identification correspond au fait que les robots peuvent être considérés comme une espèce à part entière. En ce sens, il est nécessaire de créer un fichier centralisant les identités de chaque robot afin de garantir le contrôle humain sur leurs activités.

- **Définition des principes de fabrication.** Il est essentiel que, comme l'a affirmé Bruno Maisonnier, P-DG d'Aldebaran et créateur du robot Nao, que les robots soient conçus « éthiques *by design* », des robots éthiques par nature. Pour cela, il faudrait notamment produire des chartes éthiques qui seraient imposées aux industries pour garantir cette fabrication éthique *by design*. Ces chartes éthiques pourraient par exemple comporter l'intégration dans le code source des robots les trois lois de la robotique¹ formulées par Isaac Asimov dans *Cercles*

¹ Un robot ne peut porter atteinte à un être humain, ni, en restant passif, permettre qu'un être humain soit exposé au danger ; un robot doit obéir aux ordres qui lui sont donnés par un être

vicieux (1942) et en les complétant, ces trois lois n'étant pas suffisantes. Si ces trois lois de la robotique comprennent de grandes limites, il est néanmoins nécessaire de s'en servir comme règles d'orientation.

- **Un droit des robots mondial.** Il serait faux de penser que, du fait de son caractère mondial, tout encadrement juridique national du développement industriel des robots et de l'intégration des robots dans la société serait inutile et inefficace. Si un droit international des robots peut annuler ce type d'effet, un droit national des robots pourrait aussi bien devenir un standard mondial, à l'image des lois sur l'informatique et des lois sur la protection des logiciels.
- **Certification des produits robotiques.** Il est possible d'envisager deux types de certifications : la certification par un contrôle de vérification des capacités et exerçant un contrôle au travers d'une commission, à l'image de Bureau Veritas par exemple ; ou une certification par déclaration. Il sera impératif de certifier les robots, et plus particulièrement les robots humanoïdes.
- **Documentation des algorithmes (*accountability of algorithms*).** Tous les algorithmes doivent être documentés et disponibles au contrôle.
- **Traçabilité des robots.** Assurer une traçabilité des robots est cruciale, afin qu'ils n'aient pas une complète autonomie. Les robots ne doivent pas pouvoir agir de manière anonyme ; il faut tracer leurs actions avant, pendant et après l'exécution de ces actions. La conservation des données tracées pourrait se limiter à une semaine, dans un souci de protection des données personnelles (car les robots ne vivent pas seuls). La traçabilité doit permettre d'identifier les erreurs et dangers (liés à la circulation de voitures autonomes notamment) tout en assurant la protection des données personnelles.
- **Responsabilité des robots.** Il est important de distinguer la responsabilité vis-à-vis de la victime et la chaîne de responsabilité. Concernant la responsabilité vis-à-vis de la victime, il faut appliquer une responsabilité sans faute aux robots. Pour une fonction donnée, le robot est supérieur ; mais le robot est inférieur sur le plan multifonctionnel. Les robots entraîneront moins d'erreurs que les humains sur une fonction donnée. De fait, le robot assume une responsabilité dommage vis-à-vis de la victime. La législation du Nevada qui, sur le plan de la régulation des voitures autonomes *NRS 482A* notamment, est la plus avancée au monde, dispose qu'est responsable la plateforme d'autonomie ; il est possible d'aller plus loin en affirmant la responsabilité du robot. L'humain est responsable car

humain, sauf si de tels ordres entrent en conflit avec la première loi ; un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la première ou la deuxième loi.

il est libre ; la responsabilité vient de la liberté. Les personnes morales, elles-mêmes libres, se sont vu attribuer une responsabilité. De fait, il faut introduire dans la chaîne de responsabilité une responsabilité singulière pour les robots.

- **Modification de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.** La modification de la loi informatique et liberté est nécessaire, notamment concernant les compagnons personnels. La loi n° 78-17 et le règlement (UE) 2016/679 excluent les activités purement personnelles ou domestiques : ainsi, un robot domestique n'est pas soumis aux dispositions de ces deux textes. Il est donc nécessaire de modifier ces textes afin de prendre en compte la robotique personnelle.
- **Création d'un droit à l'intimité numérique.** Il existe une co-intimité entre le robot compagnon et son utilisateur qu'il faut protéger.
- **Création d'un archivage des cerveaux robotiques.** Il serait important d'archiver les cerveaux des robots, sur le modèle des archivages bibliothécaires.
- **Droit à la transparence des algorithmes.** Il est essentiel, non pas de comprendre l'algorithme, mais de comprendre ce qu'il fait faire au robot. Le robot ne doit pas pouvoir agir de manière opaque.
- **La loyauté des robots.** Le robot ne doit pas se comporter de manière déloyale vis-à-vis de son utilisateur.
- **Création d'un droit à la compréhension.** Le droit d'information et d'accès aux algorithmes est déjà prévu par la loi de 1978 ; il est cependant nécessaire de la renforcer par l'instauration d'un droit à la compréhension.
- **Personnalité juridique robot.** Il est nécessaire de donner une personnalité juridique aux robots, afin qu'ils puissent assumer leurs responsabilités devant la loi. La création d'une personnalité juridique robot est essentielle. Les robots jouissant d'une certaine autonomie, le droit des objets ne pourra pas être appliqué.
- **Création d'un commissariat aux algorithmes.** La création d'un commissariat aux algorithmes permettra la protection des données et des règles de fabrication et d'utilisation imposées aux industries.
- **Création d'un contrat-type d'assurance pour les robots.**

- **Instauration d'une politique de label et de normes.** L'instauration d'une politique de label et de normes permettrait aux industriels de développer leur activité en minimisant les contraintes.
- **Création d'un comité d'éthique.**

2. M. Henri Verdier, directeur interministériel du numérique, ancien entrepreneur et spécialiste du numérique

La plus grande valeur dans l'économie numérique est l'intelligence partagée des individus, l'intelligence de la multitude. Les « GAFAMI »¹, ainsi que les autres entreprises majeures du numérique telles que Uber, parce qu'ils maîtrisent les stratégies de plateformes, suscitent, stimulent, concentrent et captent à leur profit l'intelligence de la multitude. **L'économie numérique croît grâce à la valeur de la multitude.**

Il n'y a pas nécessairement de frontière nette entre l'intelligence artificielle « forte », qui aurait un comportement quasi-humain, **et les systèmes experts** qui s'apparentent davantage à de l'intelligence artificielle « faible » et qui se nourrit de l'apprentissage profond (*deep learning*) et des mégadonnées (*big data*). Les mutations actuelles relèvent davantage de la progression d'une intelligence systémique à tous les échelons. **Il apparaît plus important d'envisager les évolutions de l'intelligence artificielle comme un continuum.** Il y aura peut-être quelques pointes extrêmes dans l'avancée technologique de l'intelligence artificielle, comme celle d'un ordinateur battant l'intelligence humaine aux échecs, mais il est encore plus important de réaliser que de nombreux processus seront informatisés. Les petites forces seront peut-être plus importantes que les prouesses médiatiques de l'intelligence artificielle. Par analogie, le fait que l'essor prometteur de véhicules autonomes soit imminent n'empêche pas le fait que, chaque année, les constructeurs automobiles rajoutent des fonctions d'autonomie dans les voitures.

Il existe de nombreuses inquiétudes s'exprimant autour du développement de l'intelligence artificielle. Au-delà du fait qu'il y a une grande concentration des puissances des capacités, **il n'est pas certain que l'humain puisse contrôler de manière complète les dynamiques internes d'une intelligence artificielle.** Les procédures de régulation par la norme ne seront pas suffisantes pour contrôler les actions d'une intelligence artificielle. Quelques sécurités ne permettront pas non plus d'empêcher une intelligence artificielle de s'émanciper du contrôle humain. Une intelligence artificielle peut être auto-apprenante et innovante, et il faut donc être très vigilant dans la manière dont elle sera contrôlée.

¹ Acronyme désignant les entreprises de l'économie numérique nord-américaines les plus importantes en termes de chiffre d'affaire : Google, Apple, Facebook, Amazon, Microsoft et IBM.

Du fait de l'intelligence distribuée, des codes sources ouverts (*open source*), et de la logique des organisations de l'écosystème numérique, **l'intervention de l'État dans le développement technologique et industriel de l'intelligence artificielle ne peut se faire au travers d'une planification d'objectifs, mais avec la stimulation de l'innovation.** La Silicon Valley ne peut être rattrapée grâce à des décrets.

Il manque la diffusion d'une pensée éthique. L'avènement de systèmes d'intelligence artificielle va amener des questions neuves, et les solutions proposées ne font pas nécessairement consensus. Par exemple, le développement des voitures autonomes soulève la question du comportement adopté par le système informatique pilotant la voiture lorsque celui-ci sera confronté à un danger imminent d'accident mortel soit pour son passager, soit pour un usager plus vulnérable. Pour le moment, les nombreuses solutions proposées en termes de responsabilité d'usage et d'éthique n'obtiennent pas le consensus. En cela, une pensée scientifique, et aussi une pensée sociale, sont nécessaires pour doter le développement de l'intelligence artificielle d'un cadre éthique. L'efficacité n'est pas une éthique ; si un algorithme remplace et améliore le travail humain, il sera essentiel de vérifier qu'il ne crée par des effets indirects d'exclusion. **Étant donné que l'intelligence artificielle s'invente actuellement dans un modèle économique simple, il n'y a pas de cas de conscience car il ne s'agit que de l'optimisation de techniques commerciales.** Le rôle de l'État, pour sa part, est de constamment mettre en balance l'égalité, l'équité, la justice, la croissance et le chômage ; à l'heure actuelle, les inventeurs d'algorithmes ne sont pas formés à prendre de tels facteurs en considération dans la construction de leurs algorithmes.

3. M. Laurent Alexandre, président de DNA Vision, fondateur de Doctissimo, chirurgien-urologue

Le mouvement actuel en matière d'intelligence artificielle est que **l'on se dirige vers un duopole entre les géants chinois et nord-américains.** Il n'existe pas en France et en Europe d'acteurs de l'industrie du logiciel à la hauteur des géants chinois et nord-américains. **La France et l'Europe ont des jeunes pousses prometteuses, un tissu industriel traditionnel et des chercheurs performants, mais elles sont reléguées au rang de consommateur et non de producteur de technologies.**

La France et l'Europe sont des « colonisées numériques », parce qu'elles ne produisent pas de richesses à partir du numérique, privilégiant une approche appliquée à l'informatique, c'est-à-dire une approche de vente. Le véritable enjeu concernant la révolution numérique et l'émergence de l'intelligence artificielle n'est pas le code, mais la donnée. L'apprentissage d'une intelligence artificielle nécessite beaucoup de données, mais peu de

logiciels. C'est en cela que **la France est en train de manquer le virage de l'intelligence artificielle** : la France possède les codeurs, mais elle n'a pas les données. Ce sont les géants du numérique et de l'intelligence artificielle, tout particulièrement aux États-Unis d'Amérique, qui détiennent les données. L'ensemble des données que tout utilisateur de services numériques connectés fournit quotidiennement nourrit et éduque gratuitement les intelligences artificielles des industries numériques chinoises et nord-américaines qui se trouvent en situation oligopolistique.

L'apprentissage du codage, notamment par les enfants, est une chose positive ; cependant, dans quinze ans, l'intelligence artificielle codera mieux et plus vite que les codeurs de niveau moyen. Il est important de donner une culture générale du code, mais **le codage humain médiocre ne résistera pas à l'intelligence artificielle au-delà de 2025.**

Il est essentiel de mettre en place des plate-formistes de l'intelligence artificielle en Europe. L'Europe doit retrouver de l'oxygène entre les géants numériques chinois et nord-américains. Il est donc vital qu'apparaissent des plate-formistes européens, des acteurs qui produiront et recueilleront de gros volumes de données. Il serait trompeur de raisonner par branche concernant le développement de l'intelligence artificielle ; **il ne faut pas « verticaliser » les approches de l'intelligence artificielle selon les secteurs, mais les « horizontaliser ».**

En ce sens, **il est impératif d'unifier le droit de l'information et la jurisprudence des organismes de CNIL européens**, qui ne doivent pas exiger un objectif et imposer des limites au traitement des données. **Un algorithme ne peut pas être traçable car il change en continu au fil du traitement des données** ; un choix doit donc être effectué : **l'émergence d'une industrie européenne de l'intelligence artificielle sera toujours bridée par la volonté de traçabilité des algorithmes.** La régulation trop importante peut étouffer dans l'œuf l'émergence d'une industrie européenne de l'intelligence artificielle et laisser un pouvoir oligopolistique dans les mains des industries chinoises et nord-américaines.

L'avènement de l'intelligence artificielle pourra entraîner des problèmes de requalification des industries dont la tâche sera remplacée par l'intelligence artificielle. Le populisme va croître car il n'existe actuellement pas de solutions et perspectives offertes aux actifs peu ou moyennement qualifiés, menacés par l'intelligence artificielle.

Il est essentiel de réorienter l'appareil productif rapidement. Pour l'instant, les systèmes d'intelligence artificielle sont des intelligences artificielles faibles qui ne sont pas spontanément complémentaires de l'humain ; le combat social sera le combat pour la complémentarité. Cependant, la complémentarité ne sera que temporaire : l'intelligence artificielle pourra remplacer, se substituer à l'humain quand l'apprentissage profond (*deep learning*) sera puissant et produira des intelligences artificielles fortes et conscientes d'elles-mêmes. **L'intelligence artificielle pourra se**

substituer à l'humain lorsqu'elle aura intégré la reconnaissance de formes (*pattern recognition*), mais n'aura pas de sens commun tant qu'elle ne sera pas forte et consciente d'elle-même. L'intelligence artificielle représente un choc technologique récent et explosif auquel la société n'est pas préparée.

L'Europe ne pourra percer sur le marché des « GAFAMI » et des industries numériques chinoises dans l'état actuel des choses ; sa seule chance est d'identifier le « coup d'après » afin d'être en mesure de concurrencer les États-Unis et la Chine au niveau de l'économie numérique. Cependant, **le monde politique est immature concernant ces questions.** Il est en retard et dans le déni technologique, bien qu'il soit dépendant du numérique. Dans cette perspective, l'avènement du populisme pourrait être lié à l'essor de l'intelligence artificielle, car **l'enjeu crucial sera de réorienter le tissu corporatif et industriel alors même que l'intelligence artificielle ne concernera pas toutes les branches d'activités en même temps.**

4. M. Pierre-Yves Oudeyer, directeur de recherche Inria, directeur du laboratoire Flowers, président du comité technique des systèmes cognitifs et développementaux de l'IEEE (Institut des ingénieurs électriciens et électroniciens)

Je suis directeur du laboratoire *Flowers*, et je travaille sur un domaine qui s'appelle la **robotique développementale et sociale**. Nous visons les machines et les algorithmes comme des outils pour construire des modèles d'apprentissage chez l'humain. On utilise les algorithmes pour comprendre l'évolution cognitive des enfants. J'ai auparavant travaillé dans l'entreprise Sony, au sein d'un laboratoire de recherche fondamentale.

Il y a eu des avancées ciblées dans un sous-domaine de l'intelligence artificielle, qui s'appelle l'apprentissage automatique (*deep learning*). Ces avancées techniques récentes sont aussi associées à un champ original dans la recherche sur l'intelligence artificielle : ces avancées sont développées de manière massive, un peu comme un programme Apollo, à l'intérieur d'un très petit nombre de grandes entreprises du *web* (Google, Twitter, Facebook, Microsoft, IBM et Apple). Ces grandes entreprises ont construit ces cinq dernières années des laboratoires de recherche en intelligence artificielle de taille bien supérieure à tous les laboratoires académiques du monde, et dans lesquels ils aspirent la plupart des cerveaux des meilleures universités du monde. Dans ces laboratoires de recherche, ils font des avancées fondamentales très vite, mais ils sont aussi capables de les appliquer extrêmement rapidement dans des applications utilisées au quotidien par leurs consommateurs.

Les techniques de développement de l'intelligence artificielle ne sont pas si nouvelles ; ces techniques font de la reconnaissance de forme. La reconnaissance de forme peut se faire par de l'apprentissage supervisé, qui

existe depuis très longtemps : étant donné une base de données avec d'un côté des images et de l'autre des étiquettes, **les algorithmes vont trouver des régularités dans des bases de données pour, dans le futur, être capable de deviner une étiquette.** Récemment, les chercheurs sont parvenus à faire marcher les algorithmes bien mieux qu'avant, tellement mieux que pour un certain nombre de tailles de reconnaissances, la machine est meilleure que l'humain.

La machine est meilleure que les humains, en ce qu'elle est plus discriminante, plus précise et plus efficace ou plus rapide que l'humain dans des cas particuliers. Par exemple, dans le monde médical, dans lequel les médecins doivent identifier des maladies à partir d'images, ou dans le domaine de l'astronomie, dans lesquels les astronomes identifient des structures de galaxies à partir d'images de télescopes, les algorithmes vont être capables, pour certains types d'images, de reconnaître des motifs plus rapidement et avec plus de précision que les humains, y compris les experts.

Dans un certain nombre d'applications, la combinaison de l'expertise humaine et de la machine est bien plus efficace que la machine seule ou que l'humain seul. D'un point de vue scientifique, il n'y a rien de nouveau, toutes les idées qui sont utilisées aujourd'hui étaient déjà présentes il y a une trentaine d'années. Il y a cependant deux choses nouvelles. D'une part, nous avons aujourd'hui à disposition de grandes bases de données avec les étiquettes, ce qui permet aux algorithmes de repérer des régularités qu'ils ne pouvaient pas repérer avec de plus petites bases. Ces bases de données sont le résultat de l'activité commerciale de ces grandes entreprises qui les construisent, et ont les moyens de les construire et de les entretenir. D'autre part, pour que les algorithmes repèrent des régularités, **il faut faire des calculs énormes pour pouvoir traiter toutes ces bases de données.** Jusqu'il y a peu de temps, les méthodes de calcul nécessaires n'étaient pas disponibles. Ce sont les organisations privées qui ont aujourd'hui cette puissance de calcul permettant de traiter l'ensemble des données de la base. L'utilisation de leurs bases de données leur permet de construire des algorithmes de reconnaissance de forme.

Pour ces entreprises, la valeur ajoutée ne se trouve pas nécessairement dans les algorithmes qu'elles ont mis au point, mais dans les bases de données sur lesquelles elles entraînent ces algorithmes. C'est ce qui permet à ces entreprises de publier un certain nombre de leurs algorithmes de manière ouverte (*open source*), parce que les données ont une application commerciale, et donc une valeur commerciale haute. Ces données ont beaucoup de valeur, et elles se monnaient.

Cependant, si ces systèmes sont puissants, ils ont des limites. Il est possible de faire de la reconnaissance de forme performante à partir de bases de données et des applications pertinentes, mais on est encore loin de l'intelligence humaine. **L'intelligence artificielle est encore très « stupide » par beaucoup d'aspects.**

Cette « stupidité » de l'intelligence artificielle peut être constatée notamment sur la reconnaissance d'image. En prenant deux images représentant toutes deux un camion avec quelques pixels de différence, l'œil humain ne percevra pas la différence infime entre ces deux images et reconnaîtra une même forme de camion. Cependant, pour l'algorithme, la première image reconnaîtra qu'il y a 99 % de chances pour que ce soit bien un camion ; mais sur la deuxième image, il sera certain à 99 % de chances que c'est autre chose.

Cela signifie que ces algorithmes de reconnaissance ont des taux de réussite de 99,5 % sur la reconnaissance de forme, mais que les 0,5 % d'erreur sont **des erreurs de sens commun et d'intuition énormes, qui sont incompréhensibles pour les humains**. Les humains, grâce à leur intuition et à leur connaissance générale du monde acquise au cours de leur développement, sont éduqués pour ne pas faire ces erreurs stupides.

Les algorithmes générateurs des images que doivent reconnaître la machine essaient de prendre en défaut les algorithmes de *deep learning*. Ces algorithmes de *deep learning*, lorsqu'ils sont fixés dans les applications dans le monde réel, peuvent être la cible de *hackers* qui vont développer des *stimuli* qui seront présentés à ces machines afin de mal orienter la décision à prendre dans la tâche qu'elles effectuent. Ce type d'erreur peut notamment être retrouvé dans le cas des voitures autonomes, où les algorithmes de reconnaissance de forme sont utilisés afin de prévenir les accidents. Un algorithme pilotant un véhicule autonome peut en effet être amené à devoir prendre des décisions rapides après avoir reconnu la présence d'un danger sur la route, comme, par exemple, des enfants accourant subitement sur la chaussée. Cependant, la présence d'un panneau publicitaire en bord de chaussée présentant des images sera interprétée comme un piéton traversant la rue et pourra diriger de fait le véhicule contre un mur afin de sauver un maximum de vies. Ce type de méthode statistique ne permet pas de garantir que des erreurs énormes ne seront pas faites. On ne peut donc garantir que des erreurs graves ne soient pas faites du fait des **confusions de représentation et d'interprétation des formes et images** par une intelligence artificielle.

Lorsque le législateur essayera ces technologies, il devra effectuer un changement de paradigme pour essayer de se faire une idée de leur utilité ou non. Auparavant, dans le domaine des transports, les technologies étaient présentes pour aider les gens, et notamment dans le domaine de l'aviation dans lesquels cela fait bien longtemps qu'il y a des systèmes automatisés ; cependant, l'introduction de systèmes automatisés était soumise à une certification : on pouvait prouver que ce système ne ferait pas d'erreur. Aujourd'hui, il existe de nombreux autres algorithmes d'aide à la décision humaine dans **des applications qui n'ont pas une culture de la sécurité** comme c'est le cas dans le domaine de l'aviation. C'est ainsi le cas dans le domaine de la voiture, où il n'y a pas du tout cette culture de la sécurité et où ces systèmes d'information automatisés sont, *de facto*, déjà utilisés. Une

réflexion similaire peut se faire dans le domaine de la médecine : en médecine, lorsqu'il est constaté qu'un médicament fonctionne sur 99 % des patients, ce même médicament peut avoir de graves effets sur les 1 % de patients restant.

Aujourd'hui, on ne réfléchit plus nécessairement en termes de sécurité, mais en termes de bénéfice/coût. Ainsi, pour le cas du médicament, l'interprétation du risque en fonction de la balance bénéfice/coût se base sur le calcul statistique du nombre d'années d'espérance de vies gagnées et du gain pour la sécurité sociale par rapport à ce qu'il fait perdre. **Les algorithmes ne seront donc pas dans la perspective de prévenir tout risque à l'avance, mais d'effectuer un calcul bénéfice/coût afin de garantir le meilleur résultat possible.** Cette perspective implique que, par exemple, dans le cas des voitures autonomes, un pourcentage de pertes puisse être toléré lorsque l'algorithme prend en considération la situation à laquelle il est confronté.

Dans l'état actuel des connaissances, les techniques d'apprentissage de reconnaissance de formes sont mauvaises pour apprendre en une seule fois à un algorithme. Il lui faut des séries d'exemples pour qu'un algorithme intériorise une reconnaissance de forme ou une notion.

En médecine, comme en développement de technologies, **l'incertitude occupe une grande place dans le travail du chercheur.** C'est en cela que la répétition des essais est primordiale à la vérification et à la validation du travail du chercheur.

Il existe d'autres utilisations de ces algorithmes, qui sont encore plus quotidiennes que dans les véhicules autonomes. On peut prendre l'exemple d'AlphaGo : les ingénieurs de DeepMind ont effectué un travail impressionnant avec le jeu de Go. Néanmoins, la défaite de la machine au jeu de Go montre les limites de la machine de la manière suivante : dans les deux premières parties, Lee Sedol est battu nettement, et à partir de la troisième partie, il fut capable d'adapter sa stratégie et d'équilibrer le jeu. **L'algorithme n'était cependant pas capable de s'adapter sans avoir plusieurs millions de nouveaux exemples.** C'est une différence fondamentale entre l'intelligence humaine et l'intelligence d'un système d'information. En outre, le système utilisé pour jouer contre Lee Sedol au jeu de Go ne serait pas capable de jouer aux dames contre un enfant, bien que les règles du jeu de dames soient bien plus simples à comprendre que celles du jeu de Go. Il faudra le reprogrammer entièrement pour qu'il soit capable de jouer au jeu de dames. Un algorithme pourra être très efficace pour résoudre un problème difficile, mais qui est très spécifique. **L'algorithme ne peut s'adapter aussi vite que l'intelligence humaine,** car il lui faut des millions de données et exemples pour qu'il s'adapte et comprenne de nouvelles choses.

À ma connaissance, ce type d'algorithmes est plutôt utilisé en médecine dans le cadre de l'aide au diagnostic et dans l'aide à la décision du

traitement à prescrire. Ils ne sont cependant pas encore utilisés pour les actes chirurgicaux en eux-mêmes. Certains le sont aussi en génomique, afin de repérer des structures ou des *patterns* dans l'expression des gènes.

Un autre domaine dans lequel ces algorithmes sont beaucoup utilisés est celui des filtres d'information : les moteurs de recherche, les systèmes de personnalisation des articles d'actualité, de plus en plus de sites *web*, des grands journaux, qui ne proposent pas à tous la même hiérarchie des articles. Ils choisissent la hiérarchie des titres et des articles de manière à correspondre aux préférences de l'utilisateur. Pour cela, ces algorithmes vont utiliser les clics des utilisateurs effectués auparavant comme des mesures des préférences de l'utilisateur, afin de détecter des régularités dans les articles consultés par chaque utilisateur.

Cela permet aux utilisateurs de trouver des informations pertinentes, mais le mauvais point est que ces algorithmes vont avoir tendance à « mettre la poussière sous le tapis » : de même qu'un enfant à qui on a demandé de ranger sa chambre fera disparaître la poussière sous le tapis, l'algorithme de filtrage d'information, en souhaitant satisfaire les préférences de l'utilisateur, va lui proposer uniquement les articles correspondant à ses opinions. La problématique de ce phénomène est que cela va contribuer à **enfermer les utilisateurs dans une bulle d'information** qui leur sera propre du fait des suggestions basées sur leurs préférences, et ainsi polariser les opinions. Cela peut engendrer des problèmes sociétaux et politiques fondamentaux, en particulier pour le développement des extrémismes.

La plupart des gens n'ont pas conscience qu'il y a des algorithmes derrière ces systèmes s'adaptant à leurs préférences. Il serait important que les gens soient éduqués à la compréhension des grands principes du fonctionnement de ces systèmes, pour leur permettre ainsi d'éduquer ces systèmes lorsqu'ils interagissent avec eux afin de ne pas s'enfermer dans une bulle d'information. Cependant, d'autres personnes connaissent très bien ces algorithmes et leurs effets, et vont les utiliser volontairement, comme ce fut par exemple le cas avec le logiciel de Microsoft « *Tay* » qui, à force d'indications visant à mal orienter l'apprentissage de l'algorithme par de nombreux utilisateurs, a publié sur le réseau *Twitter* des propos racistes. Cet événement est symptomatique du fait qu'il est possible de modifier le comportement de ces algorithmes qui feront à notre place des choix de filtrage d'information, ce qui soulève des questions.

Les grandes entreprises développant ces algorithmes ont pris conscience de ce problème, certaines d'entre elles ayant commencé à développer des algorithmes « *adversario* » pour contrer ces algorithmes de filtrage qui provoquent des bulles d'information. On pourrait dire que cela est une bonne chose, car cela montre que les entreprises développant ces algorithmes ont conscience du problème ; cependant, est-ce que la solution est de développer des algorithmes qui, pour votre bien, vont lutter contre d'autres algorithmes qui voulaient aussi initialement votre bien mais qui

font votre mal ? Le souci est que la décision est déléguée aux algorithmes qui vont lutter les uns contre les autres, sans vraiment poser la question.

Les systèmes d'intelligence artificielle, présents dans les voitures autonomes et dans les systèmes de filtrage d'information, fonctionnent de manière très différente des humains, en particulier parce qu'ils **sont spécialisés sur une seule tâche programmée à la main par un ingénieur**. Cependant, les enfants humains n'apprennent pas du tout de cette manière ; **l'apprentissage des enfants est d'abord dirigé par des systèmes de motivation qui sont internes à leurs cerveaux**. Les enfants ont des systèmes de motivations extrinsèques, qui vont effectuer des activités car des éléments extérieurs les y poussent, comme une récompense offerte par les parents ; et ont aussi des systèmes de motivation intrinsèques, que l'on peut associer à la curiosité, et qui vont les pousser spontanément à explorer leur environnement. Il semble que l'évolution ait doté notre cerveau de circuits spécifiques de la curiosité qui non seulement nous poussent à explorer des situations nouvelles et surprenantes, mais qui nous procurent du plaisir dans la découverte et l'apprentissage.

Dans le cadre de nos recherches, nous avons essayé de modéliser ces mécanismes, le but n'étant pas de rendre les robots plus intelligents mais de voir **comment rendre plus précis ces mécanismes d'apprentissage** en les expérimentant sur des robots, et de fournir la théorie du fonctionnement de l'enfant et de l'humain. Les mécanismes de l'apprentissage humain sont tellement complexes que l'on ne peut les comprendre par de simples matrices cérébrales ; les mécanismes d'apprentissage humain ne peuvent être totalement appliqués à l'apprentissage automatique (*machine learning*). Le robot peut tester son environnement en faisant des expériences afin d'observer des régularités, engranger de l'information en effectuant des expériences et apprendre de ses erreurs. Les expériences menées sur l'interaction du système d'apprentissage, de son système d'exploration spontanée, des propriétés de son corps et de l'environnement permettent de fournir de nouvelles hypothèses concernant les enfants. Aujourd'hui, en collaboration avec des laboratoires de psychologie et de neurosciences, il est possible de faire des expériences nouvelles avec des humains et des animaux afin d'infirmer ou de confirmer certaines hypothèses théoriques.

Cependant, les systèmes d'intelligence artificielle développés dans ce type de laboratoire visent un but très différent que les systèmes d'intelligence artificielle développés par les grandes entreprises évoquées précédemment. Il ne s'agit pas ici de résoudre des problèmes d'ingénieurs, mais d'**explorer les processus d'apprentissage et du développement**.

Il ne s'agit pas de vendre un service qui serait commercialisé, mais cela ne veut pas dire qu'il n'y a pas des applications possibles. Par exemple, le laboratoire *Flowers* travaille actuellement sur des applications dans le domaine de l'éducation. Si l'équipe de chercheurs parvient à mieux comprendre certains processus de l'apprentissage, comme ceux liés à la curiosité, cela donnera des idées d'amélioration des systèmes pédagogiques

et d'éducation. Cela pourrait également permettre de **personnaliser les parcours d'apprentissage**, et fournir des outils complémentaires au travail de l'enseignant. Un enseignant, face à trente élèves, pourra ainsi se voir fournir des outils lui permettant d'organiser la classe différemment et lui donnant les informations pour personnaliser son enseignement, ce qu'il ne peut faire mécaniquement pour trente élèves.

Ces exemples d'application sont d'ailleurs très différents des applications faites par les grandes entreprises développant des systèmes d'intelligence artificielle.

Il existe aussi des **logiciels de réadaptation ou de remobilisation** pour des enfants souffrant d'un handicap. Ce type d'application n'est pas exploité dans le laboratoire *Flowers* en particulier, mais il y a des collègues chercheurs qui se penchent sur ce type de recherche. Cependant, leurs applications visent les enfants, mais aussi les personnes âgées. Par exemple, le laboratoire *Flowers* est actuellement en train de travailler sur un projet avec l'ENST Bretagne à Brest autour des personnes âgées ayant besoin de rééducation physique et de kinésithérapie. Si la personne doit faire des mouvements physiques particuliers, le plus grand problème sera la motivation : à quel point fera-t-elle ses exercices sérieusement quand elle sera rentrée chez elle ? Avec ce projet, nous essayons de développer des technologies et des robots qui vont jouer un rôle de coach pour encourager les personnes âgées lorsque le kinésithérapeute n'est pas à côté d'elles.

Ces robots de soutien sont mis au point avec des enseignants pour notre projet porté sur l'école, et un enseignant a même été recruté afin de le mettre au cœur du développement du projet. Le laboratoire travaille très étroitement avec les écoles, les médecins et les associations de patients dans nos différents projets.

Nos collègues de Brest travaillent avec des diplômés formés à l'université dans des formations paramédicales, mais qui ne sont ni des médecins ni des kinésithérapeutes. Ces diplômés connaissent néanmoins parfaitement bien la morphologie et la biologie du corps humain, et sont capables d'interagir sur des fonctions particulières, soit pour compenser une déficience morphologique, soit pour améliorer le fonctionnement du corps pour des personnes en réadaptation fonctionnelle. Cependant, il existe également dans le domaine de l'éducation un certain nombre de spécialistes sur le terrain qui ont énormément de connaissances pratiques et qui sont là en complément des équipes éducatives.

Il est important de se demander ce que peut faire la France pour tirer parti de ces technologies. Il y a selon moi deux défis à relever. Tout d'abord **un défi sur le long terme, le défi de l'éducation**, qui est fondamental. L'éducation doit commencer dès le plus jeune âge en enseignant les principes fondamentaux de l'informatique. C'est ce qui est actuellement en train d'évoluer, avec l'intégration dans les programmes de l'école primaire de notions relatives au monde de l'informatique. Ensuite, sur le court terme,

il y a **un défi d'équilibre entre le privé et le public**. La plupart des avancées sont faites par un petit nombre de grandes entreprises disposant de laboratoires qui attirent les meilleurs chercheurs et étudiants des universités les plus prestigieuses au monde. Le défi majeur sur le court terme est donc le rééquilibrage des forces de recherche entre les laboratoires publics et les laboratoires privés.

IV. AUDITIONS DU 24 NOVEMBRE 2016

1. Mme Flora Fischer, chargée de programme de recherche au CIGREF, Club informatique des grandes entreprises françaises, pilote du groupe de travail sur l'intelligence artificielle en entreprise

Le CIGREF est une association de grandes entreprises françaises créée en 1970, qui regroupe cent quarante grandes entreprises et organismes français de tous secteurs d'activité. **Le CIGREF s'est donné comme mission de développer la capacité des grandes entreprises à intégrer et à maîtriser le numérique.** La gouvernance du CIGREF est assurée par quinze administrateurs et une petite équipe de permanents qui animent des groupes de travail, avec pour sujet l'innovation, la transformation des ressources humaines, la *blockchain*, etc. Le Cercle Intelligence Artificielle s'inscrit dans cette perspective. Initiative datant de 2015, la création du Cercle Intelligence Artificielle au CIGREF est le fruit de la concertation du président du CIGREF, M. Pascal Buffard, également président d'*Axa Technology Services*, et M^e Alain Bensoussan. **Cette initiative Cercle Intelligence Artificielle n'incluait que des entreprises « utilisatrices »,** ce qui exclut les fournisseurs et éditeurs de solutions logicielles du Cercle. Cependant, **il fut ouvert à des acteurs extérieurs - jeunes pousses, des experts, des consultants, etc. - afin d'apporter une expertise que les entreprises utilisatrices n'ont pas forcément toutes sur ce sujet.** En septembre 2016, dans le prolongement d'un colloque, le CIGREF a publié un livre blanc intitulé « *Gouvernance de l'intelligence artificielle dans les grandes entreprises : enjeux managériaux, juridiques et éthiques* » en partenariat avec le cabinet Alain Bensoussan Avocats.

Le sujet de l'intelligence artificielle demeure très prospectif : les entreprises viennent au Cercle Intelligence Artificielle principalement pour se nourrir d'informations. Le Cercle Intelligence Artificielle fournit des recommandations, qui se rattachent à trois métiers du CIGREF : l'intelligence, qui consiste à rassembler tous les savoirs des entreprises ; l'influence, qui permet au CIGREF de partager les convictions de ses entreprises membres à son écosystème ; et l'appartenance, qui repose sur les directeurs des systèmes informatiques qui doivent s'approprier ces sujets et les porter au sein de leurs entreprises.

Ce Cercle a, grâce à l'appui juridique de M^e Bensoussan, une orientation très sociétale. Le CIGREF doit également prendre en compte un enjeu managérial. **Certaines entreprises étaient plus matures que les autres sur le sujet de l'intelligence artificielle, et avaient déjà des propositions fortes.** Par ailleurs, l'équipe de prospective d'une entreprise a soulevé au niveau du management le problème du risque de la fracture d'intelligence dans les entreprises.

Dans la perspective du CIGREF, **la préoccupation majeure est celle de la transformation du management**, en premier lieu, **et ensuite celle des usages**. En ce sens, Bernard Georges¹ a affirmé que, **pour éviter la fracture intelligente qui suivrait la fracture numérique**, et pour que les entreprises ne se retrouvent pas en incapacité d'intégrer ces outils, **il faudrait que l'entreprise s'ouvre de manière plus fluide à tout ce qui est en mouvement autour d'elle et qui génère de la disruption**. Dans le schéma pyramidal et rigide des entreprises, le bas de la pyramide, en contact avec le terrain, peut s'ouvrir sur l'écosystème, tout comme le haut de la pyramide qui est en contact constant avec son environnement ; cependant, le milieu de la pyramide, qui gère beaucoup d'équipes centrales, a en quelque sorte pour habitude que les règles ne changent pas et ne peut donc pas s'ouvrir. Il y a un véritable enjeu de « fractalisation »² des organisations.

L'utilisation de l'intelligence artificielle ne peut être développée et appropriée au sein des entreprises de manière isolée ; il faut fonctionner entre communautés. Les entreprises doivent entrer soit en collaboration, soit en « co-pétition », car toutes les entreprises ont besoin et tirent un avantage des données des autres. L'intelligence artificielle repose sur *l'open source* et *l'open data* ; **il y a donc un enjeu d'ouverture qui est très compliqué à aborder dans les grandes entreprises**, pour des questions culturelles, de rigidité, et de pouvoir. Dans cette perspective, sur initiative de leur directeur des données (*CDO - Chief Data Director*), la MAIF tente de convoquer un maximum d'expertise et de disciplines différentes afin d'aborder le sujet de l'intelligence artificielle, dans l'optique de mettre en place un comité consultatif d'éthique sur la politique liée à l'intelligence artificielle.

Une entreprise a également intégré la dimension sociétale de l'intelligence artificielle dans les scénarios prospectifs élaborés sur ce sujet, notamment sur le type d'emplois qui seront impactés par l'intelligence artificielle, sur son acceptation sociétale, etc. Dans un premier scénario, les entreprises resteraient très méfiantes du fait de l'impact conséquent de l'intelligence artificielle sur l'emploi. Un autre scénario prévoit qu'une relation d'empathie se serait développée à l'égard des machines, en considérant que ce ne sont plus des contraintes comme cela pouvait être le cas avec l'informatique qui a contraint les relations lors de son introduction dans les entreprises. **Contrairement à l'informatique, l'intelligence artificielle nous amène de plus en plus vers des interfaces naturelles, qui visent à fluidifier son usage en limitant notamment le nombre d'opérations à faire (par exemple, grâce à l'échange vocal avec un *bot*) ; l'intelligence artificielle se fonde sur l'interaction.**

¹ Responsable maîtrise d'ouvrages stratégiques (innovation, systémique, prospective) à la Société Générale

² Au lieu d'une structure pyramidale, l'entreprise fractale est formée à partir d'une même structure simple et maîtrisée qui se répète à différentes échelles. Une représentation simplifiée du concept pourrait en être les poupées russes.

Le CIGREF met en avant également l'éthique *by design*, en s'inspirant du concept de *privacy by design*. L'éthique *by design* consiste à **prendre en considération le plus en amont possible de la conception, et dans le suivi d'une innovation technologique, les questions éthiques et les questions d'acceptabilité sociales**. Cela suppose d'intégrer des valeurs éthiques dans la conception même des outils numériques (applications, code informatique...).

Il existe un sous-champ de recherche de l'éthique de la technologie, la *moral ethics*, qui s'intéresse à la théorie des agents artificiels. Cette théorie consiste à chercher la manière dont peuvent être implémentés des principes éthiques chez des agents intelligents qui agiront de manière autonome. De nombreuses expériences sont menées afin de voir si une machine pouvait se comporter de manière éthique d'un point de vue humain.

L'éthique *by design* comprend deux aspects. Le premier aspect est la **mission éthique *a priori***, à effectuer en amont, car les technologies ne sont pas neutres, et **il faut donc expliciter les choix faits lors de la programmation d'un algorithme**. Le second aspect est la **mission éthique *a posteriori*, dans le suivi, avec l'apprentissage d'une intelligence artificielle**. Si la recherche parvient à inventer des systèmes d'apprentissage automatique (*machine learning*), ce vers quoi les chercheurs se dirigent petit à petit, il sera important de contrôler régulièrement comment une intelligence apprend et comment elle modélise à partir des données qu'elle a elle-même récoltées et analysées. Cependant, il n'est pas possible d'accéder à tous les niveaux de détails du processus d'apprentissage profond et automatique d'une intelligence artificielle.

L'éthique *by design* considère l'intelligence artificielle comme une substance active, à l'image d'un médicament, et implique donc de procéder de la même manière que pour la mise en vente d'un médicament, c'est-à-dire en appliquant de nombreuses séries de tests sur l'usage et l'interaction avec l'environnement de l'intelligence artificielle avant toute mise sur le marché.

Les entreprises demeurent très prudentes sur l'*open data*, davantage encore que pour l'*open source*. Les entreprises publiques ont plus de facilité à engager des initiatives d'*open data*, car la menée de ce type d'initiatives au sein d'entreprises privées sera confrontée à des questions de protection des affaires et de confidentialité de l'information patrimoniale. **Tant que les entreprises ne seront pas matures en ce qui concerne le traitement et la valorisation des données, il sera compliqué pour elles de l'être pour l'intelligence artificielle.**

Pour conclure, dans son rapport « *Gouvernance de l'intelligence artificielle dans les entreprises* », le CIGREF effectue **plusieurs recommandations** visant le développement de l'utilisation de l'intelligence artificielle dans les entreprises. Le CIGREF recommande tout d'abord d'allouer des budgets et des ressources pour l'intelligence artificielle dans les entreprises. Il est également recommandé de développer le travail de

prospective, en développant des scénarios qui permettent d'orienter les stratégies.

Viennent ensuite deux enjeux essentiels. Tout d'abord, l'enjeu opérationnel d'anticipation de la transformation des infrastructures de technologies de l'information (IT), en passant d'architectures séquentielles à des architectures parallèles. Ensuite se pose un enjeu commercial : grâce au rôle du directeur des données, l'entreprise doit réussir à faire travailler les infrastructures technologiques de base avec les métiers afin de pouvoir répondre à des enjeux commerciaux. Enfin, les entreprises doivent attirer et valoriser les talents.

2. M. Max Dauchet, professeur émérite à l'Université de Lille, président de la Commission de Réflexion sur l'Éthique de la Recherche en sciences et technologies du numérique (CERNA) d'Allistene, alliance des sciences et technologies du numérique

La CERNA ambitionnerait que les débats autour du numérique aient le même statut social que les débats sur la santé, car le numérique concerne l'avenir de l'humain de la même façon. La CERNA a été créée il y a quatre ans maintenant par Allistene (alliance des sciences et technologies du numérique) : comme le CNRS et Inria pensaient qu'il était nécessaire de mener une réflexion éthique spécifique au numérique, ils ont préféré former une commission unique au sein d'Allistene.

La CERNA est saisie ou s'autosaisit de sujets d'éthiques de la recherche en sciences et technologies du numérique. Comme la recherche est grandement liée aux usages, sachant à quelle vitesse les usages se propagent dans la société qui s'en empare, la CERNA a décidé de s'orienter vers la contribution du chercheur à tous les débats.

Actuellement, l'intelligence artificielle fait peur. C'est une peur qui est peut-être en partie orchestrée ; mais si vous voyez **les déclarations faites dans la presse**, elles sont **soit alarmistes, soit d'une opacité rassurante** dans le style « *ne vous inquiétez pas, on s'en occupe* ». Je souhaiterais vous faire partager quelques réflexions personnelles à partir de l'actualité.

Le 24 octobre 2016, MM. Martin Abadi et David G. Andersen, deux chercheurs de *Google Brain* (le programme de recherche en intelligence artificielle de Google), ont publié un article dans lequel ils expliquent comment deux réseaux de neurones apprennent à communiquer entre eux en secret d'un troisième réseau de neurones. L'expérience menée par ces deux chercheurs est une piste scientifique intéressante, et ne soulève rien d'inquiétant en termes d'intelligence artificielle. Cependant, **le fait que l'on ne parle de l'intelligence artificielle seulement lors d'accélération médiatisées de la recherche peut contribuer à susciter l'inquiétude**. La première chose à faire est de former les enseignants au numérique.

Il y a, en outre, des plateformes et des outils sur Internet qui permettent à une intelligence artificielle d'apprendre à optimiser, par exemple, le débit d'arrosage de votre jardin. Il existe des détenteurs de capitaux aux États-Unis qui se présentent ouvertement comme activistes et philanthropes tout en encourageant le développement des outils d'intelligence artificielle permettant aux gens de les prendre en main. Ces plateformes restent cependant en langue anglaise.

Il est possible d'observer, dans le cadre du développement de l'intelligence artificielle, une **double fracture : une fracture numérique, et une fracture de la langue**. Il apparaît donc essentiel de développer une plateforme ayant recours à l'intelligence artificielle en français afin de lutter contre cette double fracture.

Il existe certes un problème culturel, mais aussi un **problème statutaire dans le système français de la recherche publique**. Le système français de la recherche publique est en silo : l'interdisciplinarité est prêchée, cependant, dans une carrière, pour obtenir des promotions, il faut être très pointu dans un domaine particulier. En outre, la transmission de la connaissance et le travail d'élargissement de la connaissance en interdisciplinarité ne sont pas valorisés. Cela figure dans les fiches d'évaluations des chercheurs ; cependant, les éléments d'interdisciplinarité ne sont jamais pris en compte dans la considération d'un avancement de carrière. **L'interdisciplinarité doit être promue avant tout dans les universités.**

Le débat entourant le **développement des nanotechnologies** peut être pris en exemple : ce débat fut confié à la commission nationale du débat public, et fut très mal instruit, car ce débat s'est résumé en une succession d'affirmations et d'oppositions non scientifiques. Ce genre de débat doit avoir lieu dans les universités. En ce sens, si vous souhaitez engager un débat public sur l'intelligence artificielle, il faudrait le faire dans les universités, même si ces dernières n'auront peut-être pas l'enthousiasme qu'il faudrait.

Cependant, comme pour le débat sur les nanotechnologies, la gestion d'un important **risque industriel est à anticiper**. Il est crucial d'apprendre à **gérer le risque numérique tout comme l'est le risque industriel**, sans besoin de fantasmer. Certes, le débat est apaisé sur la question des nanotechnologies, mais cela a eu pour résultat que la France qui, en 2007, représentait 6 % de la publication scientifique sur les nanotechnologies, ne représente, en 2016, plus que 4 % de publications scientifiques sur les nanotechnologies. La Chine, pour sa part, représente 47 %, alors même qu'elle n'en représentait que 11 % il y a huit ans. De fait, il ne faut pas être paralysé par des peurs, sur le sujet des nanotechnologies comme sur le sujet de l'intelligence artificielle, car cela tournera à notre désavantage.

Les chiffres des parts de publication scientifique dans le monde sont absolus, mais il est possible de ramener ces chiffres à l'importance de la population. Dans cette perspective, le rapport entre la France et les États-Unis demeure correct ; et le rapport entre la France et la Chine demeure favorable à la France. La France représente le huitième de la production scientifique de la Chine en nanotechnologies alors que sa population est vingt fois moins nombreuse. La France n'a pas à rougir, car il y a un poids démographique à prendre en compte. La production scientifique française est excellente, mais la France ne représente que 1 % de la population planétaire, ce qui peut poser un problème en termes de relais industriels.

Il est possible d'envisager une **initiative européenne concernant l'intelligence artificielle**. Cependant, les premiers à pâtir de la montée en puissance en Europe de l'intelligence artificielle seraient les pays qui sont d'importants exportateurs en matière de mécanique de très haute précision, cette exportation se dirigeant majoritairement vers les États-Unis d'Amérique. Il y aura une rencontre entre l'intelligence artificielle et la mécanique de très haute précision (pour les robots et micro-robots) ; or, les États-Unis d'Amérique importent toutes leurs machines depuis les plus gros exportateurs européens de machines mécaniques de très haute précision, soit la Suisse, l'Allemagne, le Danemark, l'Italie, la Suède et l'Autriche. Il est également important de **prendre des initiatives d'information et de formation sur l'intelligence artificielle, mais plus largement sur les sujets du numérique, en langue française**. Pour les personnes à fort capital intellectuel dans le numérique, la confrontation à l'anglais dans leur découverte du savoir n'est pas problématique ; cependant, développer des initiatives d'information sur ces sujets en langue française permettrait d'impliquer plus largement les gens dans la compréhension du développement de ces technologies et ainsi de réduire leurs inquiétudes.

L'apprentissage d'un réseau de neurones sur de grandes bases de données d'images implique des millions de paramètres qui limitent la compréhension de ce processus. Cette complexité peut avoir pour effet de faire peur aux gens ; c'est pour cela qu'il est essentiel de les former.

La question de la gestion du risque a imprégné le monde industriel. Il y a, dans le numérique, des risques d'un autre type, parfois plus abstraits mais qui n'en nécessitent pas moins le développement d'une culture du risque numérique semblable à la culture du risque industriel : pas une culture qui vise à effrayer, mais qui vise à **mettre en place des garde-fous et des outils de précaution raisonnables**, pour que ces outils s'insèrent bien dans la société.

Il existe cependant un risque de se heurter à des difficultés institutionnelles, notamment pour l'introduction de ce type d'enseignements dans les programmes, encore, par rapport aux statuts des enseignants en fonction de leurs disciplines.

Concernant **les missions et les objectifs de la CERNA**, en couvrant tout le secteur public de la recherche dans le numérique, le principal levier est la sensibilisation et l'équipement des chercheurs pour aborder ces questions. Le but est d'intégrer davantage les chercheurs aux délibérations du pays. La CERNA pourrait disparaître au profit d'une autre structure.

La CERNA vise à se positionner comme une compétence-ressource au titre de la recherche dans le débat public, en particulier vis-à-vis de la CNIL, qui est l'ensemblier de ces questions. Actuellement, la CERNA n'est pas visible en dehors de la sphère académique. Il est souhaitable qu'une plus grande visibilité de la CERNA permette aux chercheurs de contribuer pleinement aux débats nationaux. Un certain mur de verre persiste entre les chercheurs et le reste de la société au sein du débat public. Les chercheurs sont cependant de plus en plus préparés à participer au débat public : des actions sont mises en place à destination des doctorants dans ce sens.

3. M. Cédric Sauviat, ingénieur, président de l'association française contre l'intelligence artificielle (AFCIA) et Mme Marie David, ingénieur, éditrice, membre du bureau de l'association

L'AFCIA estime que l'intelligence est la manifestation d'une puissance de calcul d'un processeur, que ce soit le cerveau, un microprocesseur ou un ordinateur. À partir du moment où est affirmé que la puissance de calcul permet d'aboutir à l'intelligence, rien n'empêche l'émergence d'une conscience artificielle : si l'on considère que la conscience est un stade d'évolution de l'intelligence, elle peut très bien émerger sur des principes artificiels.

L'accélération du progrès technique actuel rend la perspective de la création d'une intelligence artificielle douée d'une puissance de calcul comparable à celle du cerveau humain probable à l'horizon 2025-2030. Le niveau humain n'est lui-même qu'une étape dans le développement de l'intelligence artificielle : si, actuellement, l'intelligence artificielle a le niveau d'intelligence des souris, en 2025 elle aura atteint le niveau humain, et en 2030, loi d'évolution oblige, elle sera mille fois plus puissante qu'un cerveau humain. C'est ce que certains chercheurs appellent la **super-intelligence**.

L'AFCIA part également du principe que **toute la recherche en intelligence artificielle actuelle est orientée vers l'émergence de cette super-intelligence, soit délibérément, soit inconsciemment.**

L'AFCIA constate enfin que **l'émergence d'une intelligence artificielle super-intelligente, suprahumaine, ouvre nécessairement un nouveau paradigme de civilisation.** La civilisation telle que nous la connaissons actuellement est fondée sur la limite de l'intelligence humaine ; dès lors que cette limite sera franchie, un nouveau paradigme de civilisation émergera.

L'émergence de l'intelligence artificielle pose trois problèmes. Le premier problème est le **risque technologique**, qui consiste à répondre à la question : « *l'intelligence artificielle peut-elle rester sous contrôle humain ?* ». Le deuxième problème se situe au niveau des **enjeux sociaux**, des répercussions de l'intelligence artificielle sur le système socio-économique qui est le fondement de notre société. Enfin, le troisième problème relève de la **dimension anthropologique et éthique de l'intelligence artificielle**, visant à se demander quel est, dans cette évolution, le projet pour l'humanité.

Le risque technologique est largement ignoré en France, comme en attestent les déclarations régulières dans les médias de M. Jean-Gabriel Ganascia, président du Comité éthique du CNRS, qui affirme que l'émergence de l'intelligence artificielle ne pose aucun problème technologique particulier. Ce problème est néanmoins pris au sérieux par les Anglo-Saxons, en particulier par Nick Bostrom¹ qui est le philosophe le plus en pointe sur la question du risque technologique de la super-intelligence, ainsi que Bill Gates et Stuart Russell, l'un des auteurs du *best-seller* mondial de manuel d'intelligence artificielle, ce dernier considérant que « *l'intelligence artificielle est plus dangereuse que le nucléaire* ».

L'enjeu pour les Anglo-Saxons est de résoudre le problème du contrôle avant de mettre au point l'intelligence artificielle et que celle-ci n'atteigne un niveau de conscience humaine. Le parallèle avec le nucléaire est très éclairant, montrant que la découverte des effets du nucléaire, dont la première application remonte à 1945, s'est faite dans les années 1960, notamment sur les tissus humains. En 1970, le phénomène d'impulsion électromagnétique fut découvert, qui fait que lorsqu'une bombe atomique est déclenchée, une impulsion électromagnétique vient paralyser les systèmes électroniques environnants. Enfin, en 1980, le phénomène d'hiver nucléaire fut découvert, impliquant le fait que, en cas de conflit nucléaire mondial, l'humanité subirait une glaciation du fait de la formation de poussières dans l'atmosphère qui absorberaient les rayonnements du Soleil.

Il a donc fallu **cinquante ans pour maîtriser toutes les conséquences et composantes de la technologie nucléaire.** Ainsi, selon Bostrom et Russell notamment, si le même processus est appliqué pour l'intelligence artificielle, l'humanité est perdue ; il faut donc réfléchir bien en amont aux conséquences de cette invention.

L'AFCIA considère, pour sa part, que ce problème est insoluble par définition : **vouloir donner de plus en plus d'autonomie à des systèmes intelligents et, dans le même temps, chercher à les garder sous contrôle est un paradoxe.** En outre, pour l'AFCIA, résoudre le problème technique - à supposer que cela soit possible - n'empêchera jamais un détournement

¹ Directeur de l'Institut pour le futur de l'humanité à l'Université d'Oxford et Directeur du Centre stratégique de la recherche sur l'intelligence artificielle (initiative de l'Université d'Oxford et de l'Université de Cambridge)

malveillant ni une aliénation pratique, l'aliénation pratique étant la subordination consentie de l'humain à l'artificiel.

Pour les transhumanistes, pour éviter que l'Homme ne soit dominé par la machine, il faut fusionner l'Homme et la machine. La réaction de l'AFCIA est d'affirmer que le **transhumanisme consistant en une fusion homme-machine** est absurde, puisqu'il représente un suicide, **une dissolution volontaire de la personnalité dans l'inconnu** : à partir du moment où vous modifiez l'intelligence de quelque chose ou de quelqu'un, vous modifiez sa personnalité. Cette solution des transhumanistes est surtout naïve : si vous attelez ensemble un cerveau humain et une super-intelligence, il est possible de deviner que la super-intelligence, si elle réagit comme un organisme hybride dans un premier temps, prendra rapidement son autonomie, ce qui ramène à une situation de domination de la machine sur l'humain.

Concernant le niveau des enjeux sociaux, **l'humanité entre dans une ère de compétition entre les cerveaux humains et les cerveaux artificiels**. Pour une tâche déterminée, le robot coûte à peu près le même prix pour une entreprise que le travailleur. L'AFCIA considère que la théorie libérale, qui affirme que de nouveaux emplois vont émerger et remplacer ceux qui disparaissent, est caduque car cette théorie n'a jamais fait l'hypothèse que le cerveau humain était lui-même en compétition. Cette théorie libérale se fonde sur l'observation de deux mille ans d'Histoire, mais ne fait en aucun cas l'hypothèse d'une situation dans laquelle une intelligence artificielle est créée. De fait, agiter comme un dogme que de nouveaux emplois seront constitués paraît pour le moins téméraire. L'AFCIA estime que les nouveaux emplois ne se constitueront pas assez vite pour remplacer ceux qui disparaissent. Certains économistes, tels que Michael Osborne, Daniel Cohen ou Nicolas Colin, constatent que, dans vingt ans, 30 % à 40 % des emplois seront robotisés. Daniel Cohen prédit, pour sa part, que la classe moyenne sera déclassée, puisque tous les emplois de classe moyenne de supervision dans les entreprises seront confiés à des algorithmes. Pour Nicolas Colin, **les emplois nouvellement créés ne remplaceront pas en nombre suffisant les anciens**, et c'est la raison pour laquelle la France doit accélérer son développement dans l'économie numérique pour éviter que ses emplois se fassent détruire par d'autres pays ; mais il n'y aura pas d'emplois suffisants pour tout le monde.

Ce ne sont pas seulement les emplois de la classe moyenne et des ouvriers qui sont menacés, mais également des emplois d'expertise, comme les emplois de médecins ou d'avocats. L'intelligence artificielle touche des domaines de plus en plus vastes ; de fait, toutes les couches du marché du travail seront concernées par l'intelligence artificielle.

L'économie numérique, fondée sur le monopole, prépare le règne de l'intelligence artificielle. Avec la multiplication des objets connectés,

l'intelligence artificielle prépare une humanité sous surveillance (*monitoring*) permanente.

Si les emplois ne se reconstituent plus assez vite, il va falloir organiser une société sans travail. Cette société consisterait en l'émergence d'une ploutocratie rassemblant les détenteurs des entreprises numériques et ceux qui ont suffisamment de capital pour échapper au déclassement par le travail. Cette société se caractérise également par **l'aliénation des classes moyennes par le revenu minimum universel**, ainsi que par des occupations de substitution pour les masses. Enfin, **puisque l'humain n'aura plus vocation à travailler, son éducation sera négligée car il y aura beaucoup moins d'intérêt à investir dans son avenir.**

Il est possible de voir dans les prémices de cette société sans travail un lien entre le processus économique en train de se créer au niveau mondial, qui se fonde sur la disruption¹ et l'exclusion de toute une partie de la population, et la montée d'une violence sociale qui se manifeste, par exemple, dans les pays arabes qui sont le parangon de ce qui arrive aux classes modestes dans les pays occidentaux. **L'intelligence artificielle a un impact extrêmement perturbateur au niveau social.**

Enfin, l'aspect anthropologique et éthique de l'intelligence artificielle, très feutré et peu souvent évoqué, est fondamental et a motivé la fondation de l'AFCIA. L'AFCIA constate que **le développement de l'intelligence artificielle suit la logique ultime et absurde d'un système technicien**, notamment décrit dans les travaux de Jacques Ellul, fondé sur la recherche perpétuelle de la performance et du « faire mieux », aboutissant au fait que **l'Homme sera exclu dans cette recherche du mieux**. Cela est absurde, dans la mesure où la technique devrait être au service de l'Homme ; mais dans cette perspective de système technicien, la technique supprimera ultimement l'humain.

L'AFCIA considère que les mobiles philanthropiques de l'intelligence artificielle et les bienfaits qu'elle apporterait sont sujets à caution. Par exemple, lorsqu'un représentant d'une grande entreprise du numérique affirme que l'intelligence artificielle apportera beaucoup de bienfaits à la planète, il refusera cependant de prendre le risque de rejoindre la masse qui subit ce phénomène ; cela ne vaut que s'ils peuvent rester eux-mêmes en haut de la pyramide.

L'AFCIA pense que **la volonté de développement d'une intelligence artificielle de niveau humain est la négation d'un besoin psychologique fondamental de l'Homme qui est la reconnaissance par ses**

¹ Le dynamisme de l'économie numérique a des effets de destruction créatrice : une innovation et sa diffusion permet de réaliser des gains de productivité, de créer de nouvelles activités et donc des nouveaux emplois, synonymes de croissance. De fait, seuls survivront sur le marché les entrepreneurs innovants capables de restructurer leur activité, d'adapter leurs procédés, d'améliorer leur processus, de manière à bénéficier des avantages compétitifs et de devenir des leaders sur un marché.

pairs et la réalisation de soi. Ces besoins fondamentaux se retrouvent dans la célèbre pyramide de Maslow, qui a à sa base la satisfaction des besoins élémentaires de l'Homme, puis le désir de protection, et, une fois ces besoins assouvis, le besoin de reconnaissance par ses pairs, le besoin d'amour, le besoin d'estime, et enfin l'épanouissement individuel. **À partir du moment où l'Homme est remplacé par une machine et devient inutile, lui laissant plus de temps libre pour s'adonner à ses loisirs, on croit répondre à ses besoins fondamentaux alors qu'ils sont en réalité démolis.**

L'intelligence artificielle a également pour corollaire que l'on souhaite apporter à l'Homme une solution à tous ses problèmes. C'est l'avènement de l'autonomie absolue de l'Homme, qui n'a pas besoin des autres puisque tout lui est apporté par son assistant personnel, son robot et tous les systèmes robotisés à sa disposition. De fait, on fait perdre le sens de l'autre et le sens de la réciprocité à l'individu ; la société humaine est atomisée.

Au-delà de ce stade d'intervention de l'intelligence artificielle, il faut également prendre en compte le **transhumanisme**, qui consiste à modifier et augmenter l'humain afin de lui rendre service. Le transhumanisme risque, par l'augmentation de certains individus mais pas de tous, et du fait de la diversité des conditions qui seront créées par l'hybridation de l'Homme, de mettre fin à l'universalité de la condition humaine et à l'unicité de la race humaine. Cette perspective ouvre la voie à un fondement objectif du racisme, et la disparition de tous les repères éthiques qui peuvent exister dans la civilisation et l'humanité. **L'intelligence artificielle peut potentiellement signer la fin de l'éthique.**

Certains roboticiens et chercheurs affirment que l'intelligence artificielle, au même titre que l'homme, est une étape de l'évolution naturelle, qui continue son œuvre vers toujours plus d'intelligence. **L'AFCIA considère cette pensée comme une idéologie**, un système philosophique qui n'a pas de démonstration ni de scientificité, **qui est parente des idéologies totalitaires du XX^e siècle** : là où les idéologies totalitaires expliquaient la suprématie de la race supérieure ou de la classe ouvrière, cette idéologie de l'intelligence artificielle explique que le développement de l'intelligence artificielle est une évolution de la nature, et que, par conséquent, il faut l'accélérer. On retrouve dans ces idéologies une explication du passé et, en conclusion, l'accélération de l'évolution.

En reprenant l'analyse d'Hannah Arendt, dans le nazisme, il faut aider la nature à se rapprocher le plus possible de l'Homme parfait et de la race supérieure, et, par conséquent, éliminer les races inférieures ; et, dans le stalinisme, la domination de la classe ouvrière est le sens de l'Histoire, il faut de fait tout faire pour cheminer avec lui et éliminer tout ce qui s'oppose à son avènement. Il est possible de transposer cette analyse au cas de l'intelligence artificielle : l'évolution se faisant dans le sens d'une plus grande intelligence, il faut accélérer le sens de l'Histoire. Hannah Arendt a déclaré : « *l'objectif des régimes totalitaires n'est pas d'assurer la domination des*

Hommes, mais d'assurer la superfluité des Hommes. Le régime totalitaire est parvenu à ses fins lorsque l'Homme est devenu une entité négligeable superflue ». **C'est exactement le programme, dans un tout autre domaine d'idées, de l'intelligence artificielle : rendre l'Homme superflu.** Il y a une proximité des explications idéologiques à l'œuvre dans ces trois phénomènes.

Or, cette idéologie est fautive car, dans l'évolution, la conscience humaine a donné naissance à la préoccupation morale ; à partir du moment où la préoccupation morale est arrivée dans le domaine de la civilisation, toutes les idéologies d'évolution sont caduques. C'est pour cela que l'idéologie de certains chercheurs en intelligence artificielle est, aux yeux de l'AFCIA, nulle et non avenue.

Un **courant de pensée post-humaniste** peut être évoqué. Celui-ci est représenté notamment par Jean-Michel Besnier qui a tenté de décrire ce que pourrait être une éthique post-humaine : en constatant que les robots seraient considérés comme des consciences à respecter à part entière, il est possible d'imaginer un monde dans lequel hommes et robots vivent en harmonie. Selon l'AFCIA, cette vision est anthropocentrique, asimovienne, dépassée, et surtout pré-kantienne en cela que M. Besnier considère que, avec son cerveau humain, il est capable d'imaginer le fonctionnement d'un cerveau d'une intelligence artificielle et ce que pourrait être l'éthique d'une telle intelligence artificielle.

Le cerveau humain trouverait sa limitation dans sa non-perception de quelque chose dans lequel il n'y a pas d'espace et de temps ; de la même manière, un cerveau humain ne peut pas imaginer ce qu'il se passe dans un cerveau qui est mille fois plus intelligent que le nôtre. Si l'intelligence artificielle ne naît pas d'une génération spontanée, même si le cerveau humain est capable d'inventer l'intelligence artificielle, cette dernière finira par s'autoalimenter, apprendre par elle-même et deviendra de plus en plus puissante par une réaction en chaîne, échappant au contrôle et à la compréhension des humains. D'autant que l'intelligence artificielle ne s'incarne pas par des robots comme ceux décrits par Isaac Asimov ; ce sont des systèmes dynamiques qui sont plongés dans un *continuum* d'informations. Il y a une dématérialisation consubstantielle à l'intelligence artificielle qui fait que ce sont des organismes qui n'ont ni commencement ni fin.

Contrairement aux individus, auxquels on peut attribuer une limite du fait de l'individualité identifiée différenciant chacun, il n'y a pas de discontinuité entre une intelligence artificielle et une autre, puisqu'elles sont en mesure d'échanger leurs informations d'une manière qui fait que l'on peut considérer qu'il s'agit d'un seul et même système. Sur ces prémices, avec un processeur mille fois plus rapide que le cerveau humain et avec des capacités différentes de celles des êtres humains, **il est difficile de pouvoir prévoir l'éthique de ce type d'intelligence artificielle.** La seule chose qui puisse être dite est que l'on ne peut pas savoir ce que

peut être l'éthique d'une super-intelligence, et peut-être qu'il est préférable de ne pas le savoir.

Le projet de l'AFCIA se fonde sur un progrès moral et non sur un progrès technique. L'AFCIA se réfère ici à l'humanisme des Lumières. L'AFCIA considère que le fait de se réaliser par son travail au service des autres est une dignité et un droit fondamental de la personne humaine, ce qui est d'ailleurs écrit dans la Déclaration Universelle des Droits de l'Homme de 1948. Il est impératif de préserver l'unité de l'espèce humaine et l'universalité de la condition humaine, car **sans universalité de la condition humaine, il n'y a pas d'intercompréhension naturelle mutuelle et pas de possibilité de vivre ensemble ; c'est la possibilité de se comprendre qui est à la base de la civilisation**. S'il y a d'un côté des hommes, de l'autre des hommes augmentés, et enfin des robots, il n'y a pas d'intercompréhension mutuelle.

Selon l'AFCIA, le progrès technique, qui est en passe de révolutionner la civilisation, doit voir son rythme exponentiel subordonné à la stabilité et à l'harmonie sociale. La stabilité et l'harmonie des sociétés sont des enjeux bien plus importants que le progrès technique, qui ne devrait être au service que de cette stabilité. La société devrait être organisée de manière à permettre à chacun de trouver une place authentique par le service qu'il rend aux autres. Lorsque l'entreprise Uber met en place des systèmes capables de détruire toute une profession, d'un point de vue technicien, l'entreprise a raison ; mais elle ne rend pas service à la société. Il vaut mieux que l'objectif de la société soit que chacun ait une place, un rôle à jouer et quelque chose à offrir aux autres, plutôt que de chercher l'optimum et l'efficacité suprême.

Pour l'AFCIA, la course à l'intelligence artificielle, au même titre que la course à l'armement nucléaire, doit être stoppée. **La course à l'intelligence artificielle est fondée, au niveau des États, sur une volonté de puissance**, chacun devant se doter de l'intelligence la plus capable de rivaliser et de surpasser les rivaux. La course à l'intelligence artificielle est comparable à la course à l'armement nucléaire qui a eu lieu durant la guerre froide. Il est essentiel de désarmer cette course, afin d'éviter les conflits pouvant éclater entre des nations rivales.

Enfin, l'AFCIA considère que l'intelligence artificielle est, malgré tout, formidable, car il suffit de l'interdire pour redonner un sens au progrès technique, pour qu'il soit au service des humains. **Il suffit d'interdire l'intelligence artificielle pour que le progrès technique redevienne maîtrisable par l'Homme**.

L'AFCIA souhaite lancer **une campagne nationale pour offrir une base de réflexion aux citoyens**, afin de briser le discours dominant sur la disruption heureuse et les bienfaits de l'intelligence artificielle. L'AFCIA souhaite également **constituer un pôle d'opposition raisonné et non violent à l'intelligence artificielle**, car, tôt ou tard, l'opposition à l'intelligence

artificielle prendra des caractéristiques violentes. L'AFCIA entend anticiper ces mouvements sociaux de révolte afin d'offrir une alternative qui permette de canaliser le discours d'opposition à l'intelligence artificielle. L'AFCIA veut également **inverser la perception de l'intelligence artificielle auprès des jeunes**, qui peuvent être séduits par cette perspective, car ils n'en voient que l'intérêt à court terme sans prendre en considération le long terme. L'intelligence artificielle n'est pas une voie d'avenir ; c'est une voie qu'il faut combattre, et l'AFCIA veut instiller chez les jeunes le doute concernant les mécanismes qui sont à l'œuvre afin de les détourner de cette perspective. L'AFCIA cherche également à créer un réseau mondial, en traduisant leurs arguments dans toutes les langues, en particulier en anglais, en chinois et en allemand, de manière à toucher un maximum de foyers de réflexion autour de ces problématiques. Il existe un très grand tissu d'associations et mouvements de pensée techno-critiques et critiques à l'encontre de la mainmise numérique ; mais l'AFCIA est la seule association à identifier l'intelligence artificielle comme le cœur du problème, d'où le besoin d'internationalisation de son action.

Cette volonté d'internationalisation de l'action de l'AFCIA est essentielle à la réussite de son projet de constitution d'un pôle d'opposition raisonné et non violent à l'intelligence artificielle. Il n'y a aucun intérêt à garder cette réflexion nationale, qui n'aboutirait qu'à affaiblir un peu plus l'économie française, et n'aiderait pas à amplifier ce discours auprès des autres pays. Il faut adopter une réglementation progressive qui puisse avoir une portée mondiale. Il faut, pour cela, enclencher un mouvement à destination de l'opinion publique.

Il serait dans un premier temps possible d'imposer aux entreprises et laboratoires de déclarer quels sont leurs programmes de recherche en intelligence artificielle, afin que ces programmes puissent être contrôlés. Compte tenu des risques encourus, il est normal qu'il y ait un contrôle politique sur ces activités, au même titre qu'il y a un contrôle sur les activités atomiques. Ensuite, si un mouvement politique pouvait être rapidement constitué, il serait possible de se diriger vers l'interdiction mondiale des armes intelligentes, comme les drones, et accorder l'immunité aux objecteurs de conscience. Cependant, **le premier moteur de la recherche en intelligence artificielle est son application militaire** : aux États-Unis d'Amérique, la DARPA (*Defense Advanced Research Projects Agency*), qui est l'équivalent du ministère de la Défense, investit des sommes considérables dans des programmes de recherche en intelligence artificielle qui visent officiellement à appuyer l'action des militaires ; néanmoins, la perspective de nombreuses autres applications militaires possibles, comme le contrôle et l'utilisation de la pensée dans l'interface homme-machine, motive également ces investissements. Ainsi, interdire l'intelligence artificielle pourrait permettre de ralentir les investissements dans ce domaine.

En outre, de manière plus radicale, il faudrait **abolir la propriété intellectuelle sur les algorithmes d'intelligence artificielle**, ce qui couperait

le désir de bien faire chez beaucoup d'acteurs économiques. La technique dans l'intelligence artificielle est tripartite entre les algorithmes, les données et la calibration, car c'est une science très empirique. Beaucoup de programmes et algorithmes sont en *open source*, et simplement avoir accès à un programme ne permet pas d'en faire un système discriminant. En revanche, la connaissance dans la manière de calibrer et de faire converger les systèmes, en somme l'expertise telle que celle que l'on trouve notamment chez Google *DeepMind*, demeure cachée pour le moment. Faire tomber dans le domaine public l'ensemble du travail fourni par les entreprises privées dans la recherche en intelligence artificielle aura pour effet de tarir la volonté de ces acteurs économiques.

Il est également possible d'imaginer de limiter la puissance de calcul des systèmes à intelligence artificielle, en interdisant par exemple de faire fonctionner des intelligences artificielles qui dépasseraient le niveau d'intelligence d'une souris.

Enfin, pour l'ensemble des arguments développés ci-dessus, **l'AFCIA considère que la recherche en intelligence artificielle est illégitime, inacceptable, et par conséquent appelle l'interdiction pure et simple de la recherche en intelligence artificielle, ainsi que la mise à disposition à titre commercial ou à titre gracieux de systèmes à intelligence artificielle.**

Toutes ces réflexions auraient pu être menées en 1948, lorsque M. Norbert Wiener théorisa la cybernétique, ou dans les années 1960 lors de l'apparition des premiers calculateurs, car ces réflexions découlent de l'écriture d'algorithmes. Cependant, depuis quelques années, un point charnière a été atteint avec l'accélération des puissances de calcul et du développement des algorithmes. L'hiver de l'intelligence artificielle des années 1980, quand la discipline était décrédibilisée, est terminé ; la discipline est revenue en grâce du fait des avancées immenses qui ont été faites. L'AFCIA estime donc que maintenant est venu le moment d'engager ces réflexions et de lancer le débat, car la courbe d'évolution de l'intelligence artificielle est exponentielle, et dans cinq ou dix ans, il sera probablement déjà trop tard.

La plupart des grands scientifiques qui ont contribué à lancer l'intelligence artificielle comme discipline scientifique dans les années 1950, tels qu'Alan Turing ou Norbert Wiener, avaient compris que, à terme, l'intelligence artificielle ferait naître les risques et conséquences évoquées au cours de cet entretien ; mais cela demeurait théorique tant que la révolution numérique - que ces chercheurs n'ont d'ailleurs pas vécue - n'avait pas eu lieu.

Concernant la médecine et plus spécifiquement le transhumanisme, l'AFCIA distingue d'une part les interventions sur la physiologie de l'homme, et d'autre part les projets de modifications de ses capacités cognitives. Les premières sont justiciables, quant à leur évaluation éthique,

des comités de bioéthiques. En revanche, l'AFCIA s'oppose par principe aux seconds dès lors qu'ils font appel à l'intelligence artificielle, car ils sont susceptibles de modifier la nature même de la conscience avec des conséquences imprévisibles.

Il y aura toujours des exemples positifs d'applications de l'intelligence artificielle dans la société, particulièrement dans le domaine de la médecine. Cependant, en prenant l'exemple de la médecine, la machine fonctionne à partir d'un apprentissage statistique, qui comporte toujours une marge d'erreur. Si la décision humaine comporte, elle aussi, une marge d'erreur, il demeure qu'appliquer l'intelligence artificielle dans le domaine médical dessaisira l'Homme de son pouvoir de contrôle et d'expertise. En outre, les ingénieurs codant des machines en apprentissage automatique (*deep learning* ou *machine learning*) sont incapables de remonter aux causes de la décision d'un algorithme, alors qu'un médecin pourra toujours expliquer les raisons qui ont orienté son choix.

À partir du moment où l'homme considère que la machine est plus experte que lui, il va se dessaisir de son autonomie de jugement car il considérera que la machine a toujours raison, alors même qu'on ne comprend pas aujourd'hui pourquoi une machine prend telle ou telle décision. Les modèles de *deep learning* prenant en entrée (*input*) des milliards de données, il est impossible de savoir ce qu'il se déroule en sortie (*output*) avec un tel volume de données qui paramètre la décision de l'algorithme provoquant inévitablement un effet de boîte noire. En cela, la tragédie de l'humain dans une société à intelligence artificielle sera son inutilité sociale.

On ne sait pas quelle est la limite entre le vivant et l'artificiel. Il existe deux grandes écoles dans les neurosciences : les mécanistes, qui voient le cerveau comme un processeur se fondant sur des interactions électromagnétiques entre des molécules, et les « spiritualistes », qui pensent qu'il y aurait une composante qui ne se réduirait pas à un ensemble de calculs. Cependant, même en se positionnant du côté spiritualiste, les dangers présentés par le développement d'une intelligence dont le fonctionnement nous échapperait, qui ferait des choix semblables à ceux faits par une intelligence humaine tout en étant ontologiquement différente, sont concrets.

Avec les objets connectés, il y aura des analogues de systèmes autonomes dans toute la société et qui seront invisibles à nos yeux, et qui modifieront les rapports sociaux : un assureur qui refuse un contrat après que son algorithme eut mesuré le battement de cœur du client quelques jours plus tôt, un algorithme qui calcule la probabilité d'être un criminel à partir d'un visage, etc. **L'intelligence artificielle a une vocation totalitaire, elle cherche à englober tous les champs du réel.** Si l'individu a le choix ou non de mettre son sort entre les mains d'un algorithme en acceptant ou refusant de monter à bord d'une voiture autonome, il ne l'aura pas nécessairement dans de nombreux autres domaines où la présence de l'intelligence artificielle sera invisible. C'est pour cela que l'AFCIA est favorable à

l'interdiction de l'intelligence artificielle, car les individus ne pourront pas choisir tout le temps de recourir ou non à des systèmes d'intelligence artificielle.

V. AUDITIONS DU 28 NOVEMBRE 2016

1. M. Claude Berrou, professeur à Télécom Bretagne (Institut Mines-Télécom), chercheur en électronique et informatique, membre de l'Académie des sciences

*M. Claude Berrou est principalement connu pour avoir inventé les premiers codes correcteurs d'erreur quasi-optimaux, au sens de la théorie de Shannon, appelés turbo-codes¹. Ces turbo-codes sont massivement utilisés dans les troisième et quatrième générations de téléphonie mobile et dans de nombreux systèmes satellitaires. Cette invention a permis à M. Berrou d'être distingué en 2005 en recevant le prix Marconi, qui est considéré comme le prix Nobel des télécoms, et en 2007 en étant élu membre titulaire de l'Académie des sciences. À partir de 2009-2010, M. Berrou s'est concentré sur les neurosciences, en remarquant que le cerveau humain avait des propriétés de correction d'erreur : le cerveau humain est capable de corriger des approximations, des effacements, des inversions de lettres, de reconnaître un sens dans un discours, etc. Il a de fait soumis un programme de recherche à l'Agence Nationale de la Recherche sur les liens entre théorie de l'information, codage et cognition mentale. Cependant, l'ANR a refusé son programme ; M. Berrou a soumis donc son dossier au niveau européen auprès du Conseil européen de la recherche, qui a retenu son projet. M. Berrou a ainsi obtenu une bourse ERC Advanced Grant de l'ordre de deux millions d'euros de la part du Conseil européen de la recherche, ce qui lui a permis de composer une équipe de recherche interdisciplinaire réunissant un neuropsychologue, un psycholinguiste, des théoriciens de l'information, des électroniciens, des informaticiens, etc. **En cela, la composition d'équipes de recherche interdisciplinaire est indispensable pour avancer vers une véritable intelligence artificielle.***

L'intelligence artificielle est une expression très lapidaire. **L'intelligence naturelle est multiforme** mais il est quand même **possible de distinguer deux grandes catégories. La première grande catégorie d'intelligence est celle qui nous relie au monde extérieur.** Cette première catégorie d'intelligence est apparue avant la deuxième catégorie dans l'évolution, et que l'humain partage avec la plupart des espèces animales. Cette intelligence de premier type **regroupe ce qui concerne les sens, la perception, l'apprentissage, la reconnaissance et l'adaptation** - chacun est un être vivant dans un monde fluctuant, riche en détails, et cet être doit s'adapter et survivre. Aujourd'hui, dans la recherche de ce qui est appelé intelligence artificielle dans le monde de l'informatique, c'est en grande majorité l'intelligence des sens qui est recherchée. Par exemple, la voiture autonome nécessite une intelligence de la vision, du radar, etc. La voiture

¹ Le principe des turbo-codes est d'introduire une redondance dans le message afin de le rendre moins sensible aux bruits et perturbations subies lors de la transmission.

artificielle est une réussite en termes d'intelligence artificielle sur cette première catégorie de facultés permettant le lien au monde extérieur.

La seconde grande catégorie d'intelligence est propre à chaque individu. Elle englobe les facultés de l'esprit, celles qui permettent à l'humain d'imaginer, d'élaborer, d'inventer, de produire. Le seul modèle à disposition pour essayer de reproduire dans une machine des propriétés de cette intelligence créatrice est le cerveau. Le cerveau a une architecture qui est nettement différente de celle des ordinateurs classiques. Les processus de mémorisation et de traitement de l'information du cerveau s'entremêlent dans un même écheciveau de connections synaptiques qui se comptent en milliers de milliards. **Cette intelligence de deuxième type peut être appelée intelligence artificielle générale.**

Plus de 95 % de la communauté scientifique - mathématiciens, informaticiens, théoriciens de l'information - s'intéressent surtout aux sens. Cela inclut la reconnaissance de visages, de scènes, etc. Sur ce point, la France est en retard. Cependant, sur l'intelligence de deuxième type, qui ne peut être que neuro-inspirée, la France a une carte très importante à jouer, à condition que soit mise en place une recherche pluridisciplinaire. L'association de neurosciences, d'informatique, de théories de l'information et d'électronique serait très efficace. Afin d'obtenir une intelligence artificielle neuro-inspirée, il faut s'abstraire du composant. Il est nécessaire d'avoir une vision de système.

Les GAFAs sont leaders dans le domaine de l'intelligence artificielle de premier type, dans laquelle ils investissent des milliards de dollars. Concernant l'intelligence artificielle de deuxième type, neuro-inspirée, le seul modèle est le cerveau humain et celui de quelques animaux. Il existera cependant probablement une intelligence artificielle de troisième type qui consistera à reproduire au niveau informationnel l'architecture du cerveau avec plus de modules spécialisés (essentiellement les sens) et généralistes (agrégation d'informations hétérogènes) que dans le cerveau humain. Pouvoir créer une intelligence artificielle de troisième type exigera un haut niveau de compréhension du cerveau humain et de ses modules spécifiques, ses *hubs*, ses circuits de communication ; mais il devra être possible d'en mettre encore plus que dans un cerveau humain. Il existe entre deux cents et trois cents modules dans le cerveau et trente milliards de neurones dans le cortex humain ; rien n'empêchera de créer des machines qui représenteront l'équivalent de trois cents milliards ou trois mille milliards de neurones. **Ce phénomène représentera la singularité technologique.**

Les débats éthiques se construisant autour de l'intelligence artificielle sont nombreux. Cependant, il semble que ces débats soient prématurés. Si les applications d'intelligence artificielle se limitent à de l'intelligence de premier type, **le débat éthique est prématuré** : le cas typique du véhicule autonome pose certaines questions mais pas davantage que concernant le cas d'un avion de chasse. Une machine préprogrammée, qui n'apprend pas de ses expériences, qui est non apprenante, ne présente

pas plus de dangers qu'un ordinateur de bureau. **Les algorithmes étant une succession d'opérations mathématiques, il n'y a pas de question déontologique ni éthique à se poser quant à leurs résultats.** Les questions doivent davantage se poser sur les applications.

Le véritable enjeu pour la France est l'indépendance informatique, qui est tout aussi stratégique que l'indépendance énergétique ou l'indépendance militaire. Cependant, concernant l'intelligence de premier type, au vu du retard accumulé, la France risque de ne pas être acteur mais simplement consommateur. En revanche, sur l'intelligence du deuxième et du troisième type, il est possible qu'un grand programme sur l'intelligence artificielle neuro-inspirée permette à la France d'acquiescer son indépendance informatique. Sous une impulsion nationale, un programme de financement de recherche et développement, permettant d'aller jusqu'à l'application, est tout à fait possible. Les expertises en présence, la formation sont de qualité ; il demeure simplement des verrous à lever, en termes de nombre de postes, de financements difficiles à obtenir, etc.

2. M. Nicolas Cointe et Mme Fiona Berreby, chercheurs en thèse de doctorat sur l'éthique de l'intelligence artificielle

Les recherches de M. Nicolas Cointe et Mme Fiona Berreby sur la modélisation des raisonnements éthiques autour des agents autonomes artificiels dans le cadre du projet Ethicaa visent **trois objectifs : définir en quoi un agent autonome artificiel peut être éthique ; produire des représentations formelles de conflits éthiques et de leurs objets** – le conflit éthique peut être au sein d'un seul agent humain, entre un agent autonome artificiel et un utilisateur humain, ou *entre plusieurs agents* ; et, enfin, **élaborer des algorithmes d'explication et de compréhension pour les utilisateurs non experts.**

La question de l'éthique de l'intelligence artificielle peut être tout d'abord abordée sous la perspective de la représentation des conflits éthiques au sein d'un seul agent. La recherche développée par Mme Fiona Berreby consiste à modéliser des situations posant un dilemme éthique. L'une de ces situations posant un dilemme est appelée le « dilemme du tramway », qui est en deux parties.

Dans la première partie, un train roule sur une voie et l'agent a le pouvoir d'aiguiller le train. Si l'agent n'aiguille pas le train et le laisse donc poursuivre sur sa lancée, cinq personnes se trouvant sur la voie seront tuées ; cependant, si l'agent décide d'aiguiller le train, ce dernier écrasera une seule personne se trouvant sur cette voie de déviation. Selon de nombreuses études exposant la première partie du dilemme du tramway à des individus interrogés sur leur choix éthique, ces derniers estiment en majorité **qu'il est**

préférable d'adopter une décision utilitariste qui consiste à aiguiller le train afin de minimiser les pertes humaines.

Dans la seconde partie, le train se dirige toujours vers cinq personnes mais l'agent, se trouvant sur un pont au-dessus de la voie, n'a plus le pouvoir d'aiguiller le train. Le seul moyen d'empêcher le train de tuer ces cinq personnes est de faire basculer depuis le pont une personne devant le train afin de sauver les cinq personnes. Cependant, les résultats de cette partie de ces études divergent de ceux de la première partie, car dans cette configuration, **les individus interrogés refusent le raisonnement utilitariste de basculement d'une personne devant le train**, solution qui privilégie la protection d'un maximum de vie humaine.

Il existe de fait une déviation de la réflexion des individus sondés dans le dilemme du tramway. La différence de considération éthique dans les deux parties, malgré le même rapport de cinq vies humaines contre une seule, peut s'expliquer par le fait que, dans la deuxième partie du dilemme, l'agent utilise la mort d'une personne comme une cause du sauvetage des cinq personnes, alors que dans la première partie, la mort de la personne n'est pas la cause mais la conséquence du sauvetage des cinq personnes. En outre, si dans la première partie, la mort d'une personne est une conséquence indirecte, cette mort nécessite dans la deuxième partie que l'agent commette directement un acte mauvais. **C'est ici la doctrine du double effet qui modifie la perception éthique de ce dilemme** : bien que l'objectif soit de sauver un maximum de vies, le fait qu'une personne meure de la décision de l'agent trouble plus fortement cette décision lorsque l'agent doit directement agir - ici, provoquer la mort de cette personne - pour parvenir à son objectif.

L'objectif d'expériences éthiques et de dilemmes appliqués au développement de l'intelligence artificielle est, à partir de règles éthiques préétablies, de permettre à l'agent autonome artificiel de calculer à partir des faits du monde et de ses règles éthiques la réaction acceptable par rapport à un dilemme.

La question de la représentation de l'éthique peut ensuite être abordée sous l'angle des systèmes multi-agents. Un système multi-agents est un système qui comprend plusieurs entités intelligentes - aussi bien des humains que des agents autonomes artificiels, doués d'intelligence artificielle. **La discipline des systèmes multi-agents s'intéresse à la problématique de l'existence de systèmes qui comprennent des agents autonomes artificiels** et qui peuvent être homogènes - où les agents sont identiques - ou hétérogènes. De nombreux systèmes impliquent à la fois des humains et des agents autonomes artificiels, comme la défense, la finance, les transports et la santé, dans lesquels des dilemmes moraux sont en jeu.

Le problème principal abordé dans les systèmes multi-agents est celui de la coopération d'agents autonomes artificiels avec des humains ou entre agents avec des éthiques différentes. L'éthique des individus varie

selon les individus, en fonction de la culture, de l'éducation, de la religion ; les règles de valeurs et de morale peuvent différer mais coexister dans des systèmes multi-agents où interagissent des agents hétérogènes.

La morale et l'éthique sont distinctes. La morale décrit ce qui est bien ou mal dans l'absolu ; l'éthique est un moyen de réflexion qui permet de prendre des décisions justes dans un contexte précis, décisions prises selon une doctrine.

Il existe trois grandes approches formulées dans la littérature scientifique **sur la question des raisonnements éthiques des agents autonomes artificiels.**

La première approche fut apportée par un chercheur de l'armée nord-américaine, qui **proposait d'avoir une réflexion éthique avant la conception des robots, puis de restreindre le comportement des robots à uniquement des actions éthiques au regard de ce qui a été pensé avant sa conception.** C'est une forme de verrou qui est apposé sur la machine, ce qui suppose que la vérité en matière d'éthique appartienne entièrement au concepteur et interdit l'interaction éthique. Dans ce système, la machine n'a pas de représentation explicite de l'éthique ; elle n'est pas capable d'expliquer en quoi ses actions sont éthiques. Cette approche n'est pas générique : à chaque fois qu'un robot sera construit, le concepteur devra réfléchir à la dimension éthique de toutes ses actions, et si certaines sont omises, cela échappera au contrôle de la machine.

La deuxième approche est celle de l'apprentissage (*learning*), qui **implique que la machine n'a plus besoin d'être recodée à chaque fois.** Il suffit de lui montrer des cas différents afin qu'elle apprenne à agir éthiquement dans des contextes différents. Il n'y aurait cependant toujours pas de représentation explicite, la machine ne pouvant expliquer ses choix éthiques car elle n'a pas de représentation symbolique de ses connaissances. En outre, une machine qui va trop se fier au cas étudié lors de son apprentissage aura comme problème d'identifier des situations de manière trop précise par rapport à certaines autres et fera des biais de raisonnement. Au contraire, si une machine n'a pas rencontré lors de son apprentissage un cas auquel elle est confrontée, elle ne connaîtra pas la réaction appropriée.

La troisième approche est l'**approche de raisonnement symbolique.** C'est une approche fondée sur les règles, qui est générique, et permet donc à la machine de déduire de ses connaissances éthiques l'action appropriée. Cette approche permet aussi de fournir à la machine des connaissances explicites, en lui formulant des règles morales, en lui donnant des connaissances sur les valeurs et les situations. La machine sera de fait capable de déduire de manière exhaustive tout ce qu'elle peut déduire de manière logique à partir de cet ensemble. **L'objectif est de pouvoir expliquer à l'utilisateur humain pourquoi une action a été considérée comme juste, selon un ensemble de règles et de doctrines, en retraçant le processus de décision de l'algorithme.**

3. Mme Laurence Devillers, professeur d'informatique à l'Université Paris-Sorbonne et directrice de recherche du CNRS au Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (Limsi de Saclay)

Du fait du développement de l'intelligence artificielle et des systèmes experts fondés sur des algorithmes complexes, **il existe un besoin crucial de formation continue des individus sur les technologies du numérique**, à tout niveau et à tout âge. La révolution numérique est une évidence et afin d'éviter les spéculations et les peurs, mettre en œuvre une formation continue de l'individu sur les questions numériques permettrait de le guider vers l'évidence des bienfaits qu'il peut en tirer.

Dans la perspective du besoin essentiel d'éducation continue sur l'intelligence artificielle, la Maison blanche a mis en place des formations en intelligence artificielle réservées aux journalistes afin que ces derniers puissent parler notamment de l'apprentissage automatique (*machine learning*) et du codage des algorithmes sans tomber dans des facilités angoissantes dues à un manque de connaissance de ces sujets.

De manière plus générale, **les personnes portant une parole publique et étant amenées à exprimer une opinion sur ces sujets doivent être au fait des connaissances**. En ce sens, **il est essentiel de simplifier et rendre plus accessible la connaissance, ce qui implique que les scientifiques prennent publiquement la parole**. Plus l'information et la connaissance seront partagées, moins il y aura d'inquiétude à propos de l'intelligence artificielle.

Cependant, un large accès à la formation continue en matière d'intelligence artificielle n'empêche pas qu'un niveau de sécurité certain doit être assuré, notamment concernant la protection des données lors de l'utilisation de logiciels « en nuage » (*cloud computing*). **Le besoin d'éducation nécessite, au-delà de la compréhension des systèmes, la compréhension des enjeux**. Le caractère virtuel du consentement concernant la protection des données personnelles amène souvent les utilisateurs à sous-estimer l'importance du consentement qu'ils donnent en un « clic ».

Quatre leviers sont identifiés par l'Initiative pour l'éthique des systèmes autonomes de l'Institut des électriciens et électroniciens (IEEE) pour affirmer la place de l'éthique dans le développement de systèmes d'intelligence artificielle. Tout d'abord, **le premier levier est l'éducation** : les individus doivent être conscients de ce qu'est un système d'intelligence artificielle, fondé sur des algorithmes, sous ses différentes formes. **Le deuxième levier consiste à élaborer des règles morales à intégrer aux systèmes automatiques d'intelligence artificielle**. **Le troisième levier revient à se doter d'outils de vérification, d'évaluation du traitement et de mémorisation des données**. Les algorithmes ne peuvent être complètement

transparents, mais il est possible de les évaluer de différentes manières. Enfin, **le quatrième levier repose sur l'existence de règles juridiques en cas d'abus.**

L'apprentissage des machines se différencie des notions technologiques fondamentales, par rapport à la notion d'évaluation notamment. L'apprentissage d'une machine et l'enrichissement de son algorithme ne sont pas transparents. Dans le cas de l'apprentissage profond (*deep learning*), ce sont des matrices du chiffre qui structurent l'algorithme de la machine, elles ne peuvent donc être transparentes. Il faut de fait trouver de nouveaux moyens d'évaluer qui soient reproductibles. De plus, si la machine s'adapte, si elle ajoute de nouvelles données à son modèle, son modèle change. Ainsi, **si l'évaluation de la machine est reproductible malgré ses évolutions, alors un objectif de transparence est atteint.** Ce n'est pas le système qui peut être évalué de manière transparente, mais la manière dont réagit le système dans différents contextes à des temps différents d'utilisation.

En outre, **les GAFAMI ont une stratégie de récupération massive des données.** Ces données ainsi que les algorithmes et technologies utilisées par ces industries ne sont pas, pour la plupart, mises à la disposition des scientifiques, ces derniers ne constituant de fait plus des garde-fous. Il est possible d'illustrer ce point avec l'exemple de la prise en considération des enjeux éthiques. Même si les géants nord-américains de l'industrie numérique (Google, Amazon, Facebook et IBM) s'intéressent à des problématiques éthiques du développement de l'intelligence artificielle, il demeure qu'ils traitent de ces sujets entre eux, sans que la communauté scientifique n'ait de visibilité ni puisse vérifier leurs travaux.

VI. AUDITIONS DU 30 NOVEMBRE 2016

1. M. David Sadek, directeur de la recherche de l'Institut Mines-Télécom, spécialiste en intelligence artificielle

L'intelligence artificielle n'est pas un domaine nouveau ; il a connu des périodes fastes et des hivers de déclin. De nombreuses questions d'encadrement de l'intelligence artificielle sur le plan éthique se posent.

Dans cette période de renaissance de l'intelligence artificielle, certaines technologies ont atteint un niveau de maturité qui semble pouvoir tenir certaines promesses. Cependant, le potentiel de progrès en l'intelligence artificielle demeure important ; et la France y a un rôle à jouer puisque le potentiel français est garanti par des équipes de recherche de niveau mondial dans les différents domaines de l'intelligence artificielle. **L'intelligence artificielle n'est pas une discipline monolithique ; c'est un ensemble de sous-disciplines et est donc, par construction, pluridisciplinaire.**

Une confusion est régulièrement commise à propos de l'intelligence artificielle résumée à ce que l'on appelle aujourd'hui l'apprentissage profond (*deep learning*) d'un système informatique. Des avancées technologiques indéniables et remarquables ont été permises par les techniques d'apprentissage profond ; néanmoins, **il est essentiel d'éviter de réduire l'intelligence artificielle à l'apprentissage profond.** Il existe, *a minima*, deux grands pans de l'intelligence artificielle. Un premier pan serait le volet « stochastique », qui repose sur des mécanismes d'apprentissage à partir de traitements statistiques des données. Au delà des progrès sur le plan algorithmique, les avancées dans ce domaine ont été permises grâce notamment à la disponibilité de grandes masses de données et à l'augmentation de la puissance de calcul des machines.

Le second volet, que l'on peut qualifier de « cognitiviste », relève davantage de la formalisation, *a priori*, de modèles de comportements. Les travaux s'inscrivant dans ce second volet se fondent sur des modèles de raisonnement, d'inférence automatique, de représentation sémantique des connaissances, de notions d'ontologie, ... Ce second volet est partiellement éclipsé par les performances du « *deep learning* » dans le domaine de la perception. Cependant, la France dispose également d'équipes brillantes dans ce second volet, qu'il convient de valoriser à leur juste niveau. **La force de frappe de la France en matière d'intelligence artificielle sera d'autant plus importante si elle se positionne sur une hybridation des deux approches évoquées.**

Force est de constater que le domaine de l'intelligence artificielle est éclairé par la présence de grands acteurs industriels mondiaux qui mettent en place des moyens humains et matériels importants sur ce sujet. **De fait, il**

est nécessaire de voir comment le système industriel français peut produire dans ce domaine de la valeur au niveau économique et sociétal sur la base des activités de recherche et d'innovation. Un véritable écosystème de jeunes pousses du numérique tend à émerger en France, beaucoup de ces jeunes entreprises étant très performantes. Dans cette perspective, il est essentiel que les jeunes entreprises émergentes françaises ne se fassent pas happer par les grands groupes industriels nord-américains ou asiatiques, bien que bon nombre de créateurs de start-up aient cette « ambition ». **Les dispositifs de soutien à la recherche et à l'innovation se doivent d'être suffisamment incitatifs pour que le fossé entre les résultats de la recherche et la production de valeur économique soit comblé de la manière la plus efficace possible.** Dans cette perspective, il est impératif de créer une boucle vertueuse entre la recherche et l'innovation, et de développer en France une culture de l'entrepreneuriat à partir du monde académique.

Lors des deux hivers de l'intelligence artificielle, les équipes de recherche françaises se sont dispersées. **Il est essentiel de faire en sorte que ces équipes retrouvent une synergie afin que la France réacquiert sa force de frappe.** L'action de l'Association française pour l'intelligence artificielle (AFIA) pourrait participer à la redynamisation de cette synergie ; il semble que cette association s'apparente davantage à un réseau. **Les pouvoirs publics doivent favoriser l'émergence de dispositifs et de structures reflétant une dynamique, une volonté et une ambition en matière d'intelligence artificielle.** La mise en relation au sein d'un même écosystème des jeunes pousses, de certains grands groupes industriels qui font de l'intelligence artificielle un axe stratégique, et des équipes de recherche académique, sont une opportunité que la France se doit de saisir.

Les techniques d'apprentissage excellent dans le domaine de la perception, incluant la vision par ordinateur, la reconnaissance de la parole ou le traitement des images, dans lequel il existe aujourd'hui de nombreux résultats opérationnels et industriels remarquables. Cependant, concernant la cognition, la compréhension et le traitement du sens, qui caractérisent l'intelligence humaine, ces techniques d'apprentissage n'ont pas encore démontré leurs capacités. **La réalisation d'avancées concrètes de l'intelligence artificielle dans ce domaine pourra vraisemblablement émerger grâce à l'hybridation des deux approches évoquées précédemment.**

Concernant les aspects éthiques de l'intelligence artificielle, la question de l'encadrement du développement des systèmes dits intelligents est primordiale, car il peut exister des comportements délibérément malveillants. Dans cette perspective, certaines données peuvent être utilisées à des fins qui ne sont pas celles pour lesquelles elles ont été produites. En outre, le fait de ne pas maîtriser les décisions et actions des algorithmes et programmes d'intelligence artificielle peut être source de questionnements

éthiques. En effet, ces programmes peuvent avoir des comportements non souhaitables sans que ceux-ci ne soient délibérément malveillants.

À titre d'exemple est cité un cas rencontré par M. Sadek au cours de ses études qui est celui d'un programme de bataille navale utilisé par le Département de la défense des États-Unis d'Amérique pour entraîner ses amiraux. Il fut constaté que le programme, dont la consigne programmée était de maximiser la vitesse de sa flotte, coulait systématiquement les bateaux de sa flotte qui étaient touchés, car ceux-ci étaient affaiblis et donc moins rapides. La consigne d'éliminer les bateaux touchés de sa flotte n'était pas explicite dans l'algorithme implémenté ; cependant, le programme a de lui-même effectué cette inférence relativement simple.

Ce cas démontre **le manque de contrôle sur ce qui est écrit dans les programmes complexes, qui comportent beaucoup d'implicite**. En cela, la fiabilité et la conformité de fonctionnement des programmes par rapport à ce qui est strictement prévu constitue un domaine de recherche en soi et implique de nombreux questionnements éthiques. Dans le cadre de l'accélération technologique qu'entraînent les progrès de la recherche en intelligence artificielle, ces travaux de recherche éthiques sur la non-conformité de fonctionnement et donc la malveillance non délibérée potentielle de programmes devraient être renforcés et mieux valorisés, notamment en France.

2. M. Dominique Sciamma, directeur de l'école de design « Strate » à Sèvres

Aux débuts de la recherche sur l'intelligence artificielle, les enjeux se concentraient sur des problématiques de programmation. Une école de pensée, appelée le « connexionnisme », proposait une approche ascendante de l'intelligence artificielle, qui reposait sur le fait que l'intelligence est une émergence, impliquant donc que c'est l'interaction des composants qui fait émerger l'intelligence. Cette école est rapidement montée en puissance dans le monde de la recherche. Cependant, dès la fin des années 1960, un chercheur du *Massachusetts Institute of Technology*, M. Seymour Papert, a démontré que de nombreux problèmes ne pouvaient être résolus par l'approche connexionniste. Les résultats de cette démonstration ont contribué à l'arrêt des programmes de financement de l'école de pensée connexionniste. Cet échec de la théorie connexionniste a favorisé l'approche descendante de l'intelligence artificielle, celle des systèmes experts qui a permis un renouveau de la recherche dans l'intelligence artificielle jusqu'à la découverte de ses limites qui ont plongé cette recherche dans un deuxième « hiver » dans les années 1990.

L'intelligence artificielle se répand de manière massive à travers Internet et les objets. À terme, les humains seront entourés d'objets qui

seront capables de percevoir le monde, grâce à des capteurs. Ces objets auront une représentation du monde et ils pourront croiser leur représentation du monde avec les données qu'ils auront captées pour prendre des décisions. En raison de l'existence de ces trois éléments, ces objets peuvent être qualifiés de robots. Les humains seront amenés à partager leur monde avec des objets capables de prendre des décisions, et qui prendront dans certains domaines des décisions à leur place - comme le montre l'émergence des voitures autonomes par exemple. De fait, le rapport des humains aux machines sera altéré : une véritable relation entre la machine et l'humain, tous deux capables de décider et de dialoguer, se substituera au contrôle exercé par l'humain sur la machine.

Les objets robotisés, les « robjets », s'invitent dans la vie quotidienne des utilisateurs ; c'est la première fois que l'humanité se trouverait confrontée à des objets non plus définis et caractérisés par des formes mais par des comportements. Ce ne seront plus des objets de contrôle ni des extensions du corps humains, ils seront dans la position d'un majordome et vont, sur la base de certaines informations, prendre des décisions. Les êtres humains accueilleront dans leur environnement proche des objets adoptant un comportement humain, avec une forme d'empathie.

Un rapport de relation va se substituer à un rapport de contrôle : les robots seront capables d'observer, de prendre des décisions et de dialoguer avec les êtres humains. Certains objets actuels, grâce à de nombreuses technologies notamment biométriques, sont capables d'observer les réactions des êtres humains ; cependant, le rapport de relations sera bouleversé par le caractère empathique qu'auront intégré les robots dans leur rapport aux êtres humains. L'intelligence artificielle peut s'incarner et s'adapter aux êtres humains avec qui elle est en relation en fonction de leurs caractères et manières de vivre.

Cette situation nouvelle impliquera de façonner les nouvelles relations qu'entreprendront les êtres humains avec les robots. Cela relève non d'une problématique de marché ou d'une question technique d'ingénierie, mais du *design* comportemental et du *design* de situation. De fait, il est nécessaire de réunir des spécialistes de la gériatrie, du personnel d'assistance, des ingénieurs, des sociologues, des spécialistes du comportement, et des *designers* pour fournir un travail d'observation, de scénarisation, d'hypothèses et de tests sur le comportement des robots dans une situation sociale.

Au-delà des enjeux comportementaux et d'acceptation sociale de l'intelligence artificielle, notamment de sa matérialisation physique, il existe également de nombreux enjeux économiques et industriels. La France est, depuis toujours, puissante en matière de mathématiques, d'informatique et d'intelligence artificielle ; à ce titre, Inria est un lieu de production intellectuelle et d'innovation remarquable. Cependant, si les premières commercialisations de produits informatiques ne visaient qu'un public

restreint et ne représentaient donc pas un enjeu économique majeur, c'est aujourd'hui plusieurs milliards d'individus qui utilisent ce type d'objets.

En ce sens, les enjeux d'acceptation sociale et de désirabilité des objets, qui ne sont pas techniques mais économiques, sont déterminants, et reposent sur des critères d'expérience et d'imaginaire associés. Par exemple, il est possible de citer le cas de l'entreprise Apple : indépendamment des qualités techniques des produits, il existe un imaginaire associé à la marque Apple, amenant le consommateur à préférer, à qualité et à esthétique égale, un produit Apple à un produit Samsung. Ce travail profond de la relation à l'objet, de son inscription durable dans l'environnement humain et jusqu'à la désirabilité de l'objet, relève des *designers*.

De fait, il existe un enjeu industriel fort. Il est essentiel que les industriels français se mettent en position de produire des objets désirables, achetés dans le monde entier. L'industrie française a déjà démontré, à maintes reprises, sa capacité à produire des objets désirés dans le monde entier, à l'instar du succès que rencontrent les drones de l'entreprise Parrot et les objets connectés de la marque Withings. L'industrie française est remarquable en matière de *design* des produits ; l'expertise française n'est plus à démontrer.

Cependant, la capacité des entreprises françaises à transformer une expertise scientifique et technique en objets industriels pose question, car cette transformation ne repose pas uniquement sur l'excellence technique. Cette transformation qui ne relève pas du *marketing* mais du *design*, de la capacité à s'inscrire dans la situation de vie humaine et donc à anticiper les décisions des robots, est, sur le plan éthique, indispensable. L'ivresse technique doit être tempérée par une réflexion sur les impacts politiques et sociaux de ces « robjets », et sur les outils de contrôle à disposition des utilisateurs et des pouvoirs publics.

La question du droit des robots représente également un enjeu à prendre en considération. Il existe deux écoles d'intelligence artificielle sur cette question : l'école de l'intelligence artificielle forte (dont je suis un tenant), et l'école de l'intelligence artificielle faible. L'école de l'intelligence artificielle forte postule que l'Homme est capable de réaliser avec l'intelligence artificielle ce que la Nature a été capable de réaliser avec lui ; de fait, à terme, l'Homme produira des machines conscientes. Dans cette perspective, la question du droit des robots se posera en ce que ces derniers seront doués de conscience, et donc potentiellement porteurs de revendications en matière de droits fondamentaux.

L'introduction de robots doués d'intelligence artificielle forte peut amener de nombreux débats philosophiques, éthiques, et juridiques. Il est important d'envisager le type de société qui découlera de l'avènement d'une intelligence artificielle forte : les Hommes et les machines entreront-ils en compétition ? L'intelligence artificielle sera-t-elle susceptible de se substituer à l'intelligence humaine ? L'importance du débat et les inquiétudes

entourant le développement de l'intelligence artificielle se justifie par le fait que cette question touche à quelque chose de profondément humain, et nécessite une réflexion éthique, politique, juridique et sociale approfondie.

3. M. François Taddéi, directeur du Centre de recherches interdisciplinaires (Inserm, Université Paris-Descartes), biologiste

Il est important que les acteurs publics s'emparent du sujet de l'intelligence artificielle, afin de ne pas laisser les seules sociétés privées, principalement nord-américaines et chinoises, nourrir de l'intérêt pour ce thème. L'intelligence artificielle va modifier les sociétés humaines en profondeur sans que celles-ci ne le comprennent nécessairement ni ne puissent en minimiser les dérives potentielles. Il existe de fait un besoin important de formation. Par ailleurs, il est étonnant de constater que davantage de recherches sont menées sur l'intelligence artificielle et l'apprentissage automatique que sur l'intelligence humaine et l'apprentissage des enfants, ce que démontre l'importance des intérêts économiques dans les orientations de recherche.

Cependant, la combinaison entre l'intelligence humaine et l'intelligence artificielle peut permettre aux humains, au-delà de la création de machines, d'accomplir des progrès qu'ils ne sauraient atteindre seuls. Dans ce contexte, les réflexions éthiques, politiques et sociétales sont essentielles, afin que celles-ci permettent de penser l'intelligence collective entre les humains et les machines.

Devant l'accélération des progrès dans l'intelligence artificielle, les systèmes d'apprentissage profond et l'apprentissage automatique, les systèmes d'éducation, conçus au XIX^e siècle, ne sont plus en mesure de s'adapter à ces évolutions toujours plus rapides, que ce soit sur le plan de la société, du marché de l'emploi, ou des défis technologiques, scientifiques et éthiques. Si la société évolue plus vite que son système éducatif, ce système éducatif deviendra rapidement obsolète. Il apparaît ainsi essentiel d'adapter le système éducatif à cet état de fait, en apprenant aux élèves non seulement à penser sans la machine, mais également à penser avec la machine. Le système éducatif actuel met en compétition les élèves sur les savoirs d'hier ; le défi sera de construire le savoir de demain et de créer les machines qui permettront la construction de ce savoir de demain. Il est de fait nécessaire de repenser le système éducatif, dont l'évolution ne peut se décréter de manière verticale, mais plutôt de manière plus horizontale, en faisant appel à l'intelligence collective de tous les acteurs éducatifs, le ministère de l'éducation nationale, les enseignants, les chercheurs et les parents d'élèves.

En réponse aux défis de l'intelligence artificielle, il est important de mobiliser l'intelligence collective humaine et toutes les générations à tous les

niveaux. Selon Aristote, les trois piliers essentiels de la connaissance : *épistèmè* - qui a donné « science » -, qui repose sur la connaissance du monde ; *technè* - qui a donné « technologie » -, qui repose sur la manière d'agir sur le monde ; et *phronésis*, qui est l'éthique de l'action. Or, à mesure que les savoirs s'accumulent et qu'ils permettent d'agir sur le monde, il est important de s'interroger sur l'éthique des actions. Les humains ont fait croître de manière exponentielle leurs connaissances et leur capacité à agir sur le monde, mais très peu leur capacité à penser les implications individuelles et collectives de leurs choix, tant sur le court terme et le long terme, que sur les plans locaux et globaux. L'avènement de l'intelligence artificielle et de la machine impliquent une plus grande considération de l'éthique de l'action. Les humains qui créeront ces machines, qui seront douées de *technè* mais très peu de *phronésis*, devront interroger l'éthique de leur action.

4. M. Igor Carron, entrepreneur, organisateur du principal « *meet-up* » en intelligence artificielle en France intitulé « *Paris Machine Learning* »

L'évolution de l'intelligence artificielle a connu plusieurs vagues, notamment aux États-Unis d'Amérique, les périodes séparant chacune de ces vagues étant qualifiées « d'hiver de l'intelligence artificielle ». **Les hivers de l'intelligence artificielle sont les périodes où le financement public nord-américain des activités de recherche en intelligence artificielle s'est estompé drastiquement.** Les réductions conséquentes de ces financements publics qui ont entraîné ces hivers s'expliquent par le fait que les pouvoirs publics estimaient que certaines méthodes utilisées ne permettaient pas d'atteindre les objectifs promis. Si la recherche en intelligence artificielle a connu un renouveau dans les années 1980, les financements octroyés à la recherche en intelligence artificielle par la *DARPA - Defense Advanced Research Projects Agency* - se sont à nouveau asséchés au début des années 1990, les algorithmes construits présentant de nombreuses limites.

La communauté scientifique de recherche en intelligence artificielle a continué à survivre de manière dynamique. **L'avènement de Google en 1998 constitue, d'une certaine manière, un nouveau point de départ pour la recherche en intelligence artificielle.** En effet, l'entreprise, dès ses débuts, a utilisé un algorithme permettant de donner des résultats aux humains. Les moteurs de recherche existants dans les années 1990, tels que ceux de Yahoo! Portal, utilisaient des humains pour effectuer une partie des classifications. **L'émergence du moteur de recherche de Google a démontré que l'humain n'était pas indispensable pour produire des pages de résultats de recherche.** En ayant recours à des algorithmes efficaces, Google a capté les utilisateurs des autres plateformes de moteur de recherche et s'est placée en situation de quasi-monopole dont le moyen de financement

principal est la publicité ciblée, qui utilise également des algorithmes. Dès 2001, les dirigeants de Google indiquaient que les activités de moteur de recherche constituaient une forme d'intelligence artificielle.

La majorité des algorithmes développés et utilisés actuellement n'offre pas la compréhension des décisions qu'ils prennent pour les humains. Il est de fait essentiel de construire des algorithmes permettant de comprendre et décoder les décisions prises par d'autres algorithmes, quitte à découvrir pourquoi ces algorithmes ont des biais ou pourquoi ils sont discriminants.

5. M. Jill-Jënn Vie, chercheur en thèse de doctorat à l'École normale supérieure Paris-Saclay

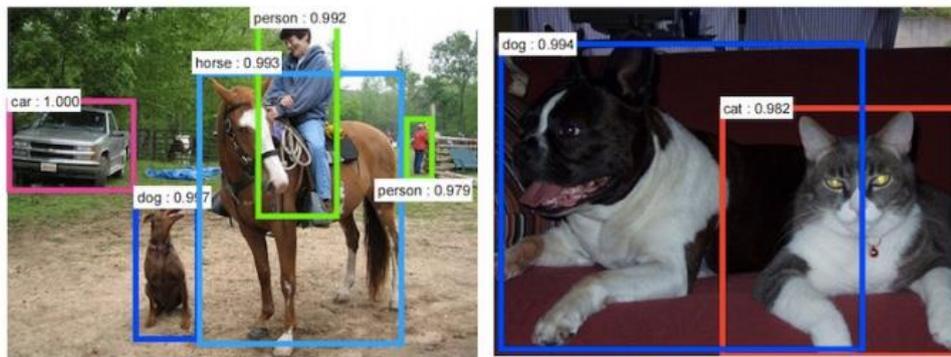
Le principe de l'intelligence artificielle consiste à inférer des faits à partir de données qui n'ont pas été explicitement programmées. Le développement d'une intelligence artificielle amène la question de l'objectif à optimiser, que ce soit le profit, l'économie d'énergie – à l'instar de l'entreprise Google *DeepMind* qui utilise son système d'apprentissage profond pour optimiser sa consommation d'énergie et réduire de 40 % la consommation d'énergie de ses centres de données –, ou encore le nombre de votes obtenus à une élection, certaines applications permettant d'optimiser l'efficacité des déplacements de campagne électorale d'un candidat.

L'utilisation de l'intelligence pour ces types d'applications a été renforcée grâce aux mécanismes d'apprentissage profond (*deep learning*). L'apprentissage profond permet entre autres à une intelligence artificielle d'effectuer de la reconnaissance d'image, de la détection d'éléments ou d'êtres vivants, de la légende automatique et de la génération d'image.

Exemple de reconnaissance d'image



Exemple de détection



Exemple de légende automatique

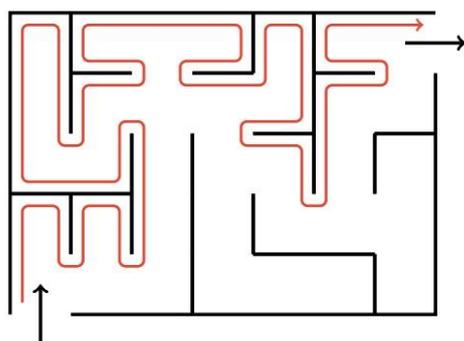


Exemple de génération d'image

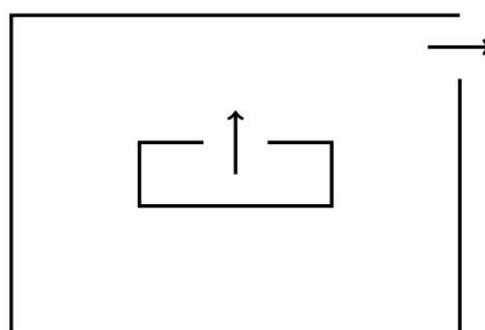


Le code source permet l'exécution de l'algorithme : il renferme l'ensemble des séries d'opérations que doit effectuer la machine dans un ordre prédéfini. La machine va compiler le code source dans un langage, qui n'est certes pas compréhensible par l'humain, mais qui lui permet d'exécuter plus rapidement les tâches qui lui sont assignées. Un exemple d'algorithme permettant la résolution d'un problème est celui dit du « parcours main gauche », qui consiste à suivre le mur à sa gauche afin de sortir d'un labyrinthe. Cependant, de nombreux concepteurs d'algorithmes cherchent à trouver des cas spécifiques pour lesquels l'algorithme ne pourra pas résoudre le problème.

Modélisation d'un « parcours main gauche »



Un labyrinthe pour lequel le parcours main gauche échoue



De nombreux jeux mathématiques sont développés sur Internet pour apprendre à programmer (par exemple un « parcours main gauche ») en assemblant des blocs : c'est le logiciel *Blockly*. Ils font fureur chez les plus jeunes (cf. code.org/frozen) et le logiciel *Scratch* utilisé pour faire découvrir l'algorithmique au collègue est basé sur le même principe.

Exemple de « Blocky game »



Lors de la programmation d'un algorithme, il est important de distinguer les erreurs de spécification et les erreurs d'implémentation. Une erreur de spécification correspond à un cas non traité ou mal traité dans le cahier des charges, ce qui a notamment provoqué l'explosion du vol inaugural du lanceur européen Ariane 5. Une erreur d'implémentation correspond à un bogue informatique lors de la programmation proprement dite de l'algorithme.

Concernant le cas des **erreurs de spécification**, à titre d'exemple, l'entreprise de commerce électronique Amazon avait mis en vente une série de T-shirts sur lesquels il était inscrit des formules comme « *Keep calm and rape a lot* » (« Restez calme et violez beaucoup »). En réalité, l'entreprise n'était que le revendeur de ces T-shirts, ces vêtements étant produits par une entreprise utilisant un algorithme qui créait de manière aléatoire les slogans inscrits. Ce produit n'avait pas été pensé, mais créé par une intelligence artificielle.

De même, certains livres vendus sur la plateforme de vente en ligne sont des ouvrages entièrement créés par des algorithmes, qui n'existent pas mais qui sont issus d'un procédé de « génération automatique de texte ». De fait, les acheteurs, pensant acquérir un ouvrage rédigé par un homme, reçoivent un livre dont le texte est automatiquement et aléatoirement produit depuis une base de données par un algorithme. L'étendue du nombre de cas similaires à ceux-ci s'explique par le fait que l'impression à la demande, et plus largement la fabrication à la demande, est une pratique régulièrement utilisée par les vendeurs présents sur Amazon, afin de rationaliser les coûts de production. Ainsi, les articles, aussi illicites soient-ils, n'existent pas physiquement avant qu'un acheteur ne les commande.

Les cas d'**erreurs d'implémentation** peuvent, pour leur part, également avoir des conséquences sur les conditions d'utilisation et le niveau de sécurité d'une machine. Il est possible de mobiliser le cas de la PlayStation 3, la console de jeux vidéo produite par la société japonaise Sony.

Afin de sécuriser l'utilisation des jeux vidéo et lutter contre le piratage, la console ne peut exécuter que des jeux bénéficiant d'une licence. La vérification de la licence repose sur un nombre secret. Durant la phase d'implémentation, les développeurs devaient choisir un nombre au hasard pour chaque nouveau jeu ; cependant, ces développeurs ont commis l'erreur d'utiliser le même nombre secret pour chaque nouveau jeu. Cette erreur fut exploitée par des hackers qui purent résoudre une équation du premier degré et trouver ce nombre secret, leur permettant de faire lire un jeu détenteur de licence sur leur console. À la suite de la publication sur Internet de ce nombre, des poursuites ont été engagées par l'entreprise, et le nombre fut considéré comme un périphérique de contournement de sécurité, ce qui rendait illégale toute publication et transmission de ce nombre.

De nombreuses plateformes de vente utilisent des algorithmes ayant recours à des formes de discrimination volontaire avec une fixation de prix dynamique. Cette technique de fixation de prix dynamique est notamment utilisée par les plateformes de vente de billets des compagnies de transports aériens et terrestres commerciaux.

Il est également possible de recourir à des pratiques de camouflage volontaire (*obfuscation*), qui consistent à ne pas ouvrir un code source de manière délibérée. Cette technique est notamment utilisée dans des applications de prédiction de crimes (*PredPol*), dans l'affectation d'un degré de menace contre les habitants lors des patrouilles de police à partir des plateformes *Beware* ou *Twitter*, ou encore dans le diagnostic de personnes à risque de pédophilie avec la plateforme *Abel Assessment*.

Tous ces systèmes boîte noire se cachent derrière le prétexte qu'une ouverture du code pourrait influencer un comportement de contournement de la part de malfaiteurs, mais leur légitimité doit être remise en question.

Néanmoins, **l'utilisation d'algorithmes peut amener à des discriminations involontaires exercées sans que cela ne soit programmé en amont** comme, par exemple, avec l'application *Street Bump*, qui vise à détecter les nids-de-poule sur la route. L'effet discriminatoire involontaire induit par l'utilisation de cette application réside dans le fait que la présence de nids-de-poule sera détectée de manière plus rapide dans les quartiers où réside une population aisée. En effet, davantage d'habitants disposent des moyens financiers pour détenir un *smartphone* capable d'exécuter cette application dans les quartiers aisés que dans les quartiers populaires.

En connaissant les règles qui caractérisent l'algorithme, on peut repérer ce type de biais et y réagir. Sinon, on ne peut pas.