

GROUPE D'EXPERTS
INDEPENDANTS DE HAUT NIVEAU SUR
L'INTELLIGENCE ARTIFICIELLE
CONSTITUE PAR LA COMMISSION EUROPEENNE EN JUIN 2018



DEFINITION DE L'IA:
PRINCIPALES CAPACITES ET DISCIPLINES

Définition élaborée aux fins des contributions du groupe

Définition de l'IA: principales capacités et disciplines scientifiques

Groupe d'experts de haut niveau sur l'intelligence artificielle

Clause de non-responsabilité et utilisation du document: La description et la définition qui suivent des capacités et domaines de recherche de l'IA simplifient à l'extrême l'état actuel de la technique. Le présent document n'a pas pour objet de définir de manière précise et exhaustive l'ensemble des techniques et capacités de l'IA, mais plutôt de proposer une synthèse de la compréhension commune de cette discipline qu'applique le groupe d'experts de haut niveau à ses contributions. Nous espérons toutefois que le présent document pourra également servir de point de départ pédagogique utile pour les non-spécialistes de l'IA, qui pourront alors s'intéresser à l'IA sur la base d'une réflexion plus large et profonde afin d'obtenir des connaissances plus précises de cette discipline et technologie.

Le GEHN IA est un groupe d'experts indépendants constitué par la Commission européenne en juin 2018.

Personne de contact Nathalie Smuha – coordinatrice du groupe d'experts de haut niveau sur l'IA
Adresse électronique CNECT-HLG-AI@ec.europa.eu

Commission européenne
B-1049 Bruxelles

Document rendu public le 14 avril 2019.

Un premier projet de ce document a été publié le 18 décembre 2018, avec le premier projet de lignes directrices du GEHN IA en matière d'éthique pour une IA digne de confiance. Il a été révisé à la lumière des commentaires reçus dans le cadre de l'Alliance européenne pour l'IA et de la consultation ouverte relative au projet de lignes directrices. Nous souhaitons remercier explicitement et chaleureusement toutes les personnes qui ont fait part de leurs commentaires sur le premier projet de ce document.

Ni la Commission européenne ni aucune personne agissant au nom de la Commission n'est responsable de l'usage qui pourrait être fait des informations données ci-après. Le contenu du présent document de travail relève de la seule responsabilité du groupe d'experts de haut niveau sur l'intelligence artificielle (GEHN IA). Bien que des membres du personnel des services de la Commission aient facilité la préparation de ce document, les avis qui y sont exprimés reflètent l'opinion du GEHN IA et ne peuvent, en aucune circonstance, être considérés comme constituant une prise de position officielle de la Commission européenne.

De plus amples informations sur le groupe d'experts de haut niveau sur l'intelligence artificielle sont disponibles en ligne (<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>).

La politique de réutilisation des documents de la Commission européenne est régie par la décision 2011/833/UE (JO L 330 du 14.12.2011, p. 39). Pour toute utilisation ou reproduction de photos ou d'autres éléments non couverts par le droit d'auteur de l'UE, l'autorisation doit être obtenue directement auprès des titulaires du droit d'auteur.

DEFINITION DE L'IA:

PRINCIPALES CAPACITES ET DISCIPLINES SCIENTIFIQUES

Nous prenons pour point de départ la définition suivante de l'intelligence artificielle (IA), telle que proposée dans la communication de la Commission européenne sur l'IA¹:

«L'intelligence artificielle (IA) désigne les systèmes qui font preuve d'un comportement intelligent en analysant leur environnement et en prenant des mesures – avec un certain degré d'autonomie – pour atteindre des objectifs spécifiques.

Les systèmes dotés d'IA peuvent être purement logiciels, agissant dans le monde virtuel (assistants vocaux, logiciels d'analyse d'images, moteurs de recherche ou systèmes de reconnaissance vocale et faciale, par exemple) mais l'IA peut aussi être intégrée dans des dispositifs matériels (robots évolués, voitures autonomes, drones ou applications de l'internet des objets, par exemple).»

Dans le présent document, nous élargissons cette définition pour clarifier certains aspects de l'IA en tant que discipline scientifique et en tant que technologie, avec pour objectif d'éviter les malentendus, de parvenir à des connaissances communes partagées de l'IA pouvant également être utilisées de manière efficace par des non-spécialistes de l'IA, et de fournir des informations utiles pouvant servir à la discussion tant sur les lignes directrices en matière d'éthique que sur les recommandations en matière de politique dans le domaine de l'IA.

1. Systèmes d'IA

Le terme «IA» comporte une référence explicite à la notion d'intelligence. Toutefois, étant donné que l'intelligence (tant chez les machines que chez les êtres humains) est un concept vague, bien qu'elle ait fait l'objet d'études approfondies de la part de psychologues, de biologistes et de neuroscientifiques, les chercheurs du domaine de l'IA utilisent principalement la notion de rationalité. Cette notion désigne la capacité de choisir la meilleure action à prendre pour parvenir à un objectif donné, sur la base de certains critères à optimiser et des ressources disponibles. Bien sûr, la rationalité n'est pas l'unique ingrédient du concept de l'intelligence, mais elle en est un élément important.

Dans le présent document, le terme «système d'IA» désigne toute composante, tout logiciel et/ou matériel fondés sur l'IA. En effet, plutôt que des systèmes autonomes, les systèmes d'IA sont en général *intégrés* en tant qu'éléments de systèmes plus vastes.

Un système d'IA est donc avant tout rationnel, selon un des manuels de l'IA les plus utilisés.² Comment un système d'IA peut-il donc faire preuve de rationalité? Comme souligné dans la première phrase de la définition de travail de l'IA ci-dessus, un système d'IA fait preuve de rationalité en percevant l'environnement dans lequel il est plongé à travers des capteurs, en collectant et interprétant ainsi des données, en développant un raisonnement sur ce qui est perçu, ou en traitant les informations obtenues à partir de ces données, et en décidant quelle action est la plus adaptée pour agir ensuite en conséquence, au moyen d'actionneurs, ce qui a pour effet éventuel de modifier l'environnement. Les systèmes d'IA peuvent utiliser des règles symboliques ou apprendre un modèle numérique. En outre, ils peuvent également adapter leur comportement en analysant la manière dont l'environnement est affecté par leurs actions antérieures. La figure 1 illustre ce fonctionnement des systèmes d'IA.

¹ Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions – L'intelligence artificielle pour l'Europe, Bruxelles, 25.4.2018, COM(2018) 237 final.

² "Artificial Intelligence: A Modern Approach", S. Russell and P. Norvig, Prentice Hall, 3rd edition, 2009. «Artificial Intelligence: A Modern Approach», S. Russell et P. Norvig, Prentice Hall, 3e édition, 2009.

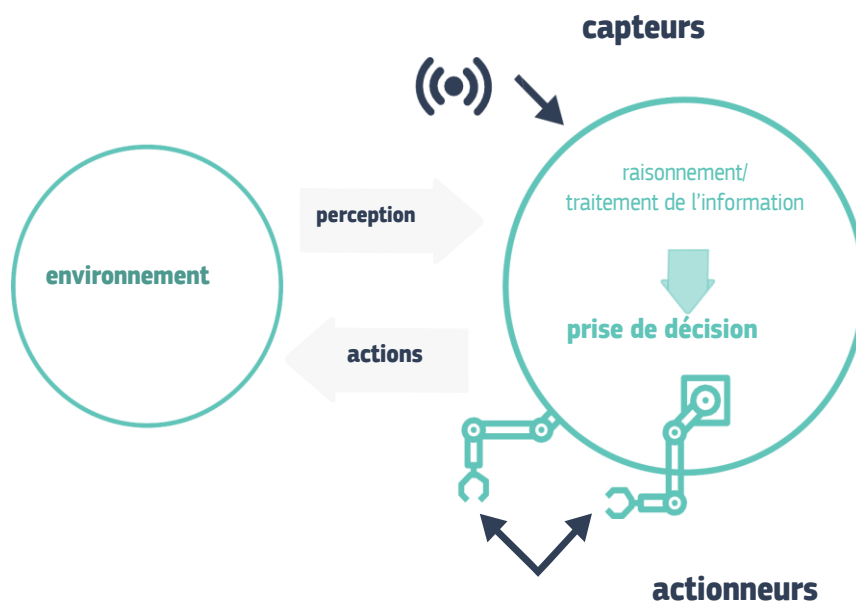


Figure 1: Représentation schématique d'un système d'IA.

Capteurs et perception. À la figure 1, les capteurs du système sont représentés par le symbole du wifi. Dans la pratique, il peut s'agir de caméras, de micros, d'un clavier, d'un site web, ou d'autres dispositifs de saisie, ainsi que de capteurs de valeurs physiques (telles que température, pression, distance, force/couple, capteurs tactiles). En règle générale, il est nécessaire de doter le système d'IA de capteurs adaptés pour percevoir les données présentes dans l'environnement qui sont pertinentes par rapport à l'objectif assigné au système d'IA par son concepteur humain. Par exemple, pour mettre au point un système d'IA qui nettoie automatiquement le sol d'une pièce lorsque celui-ci est sale, les capteurs pourraient comprendre des appareils photo qui prendront des photos du sol.

En ce qui concerne les données collectées, il est souvent utile de faire la distinction entre les données structurées et les données non structurées. Les *données structurées* sont les données qui sont organisées en fonction de modèles prédéfinis (comme dans une base de données relationnelle), tandis que les *données non structurées* ne sont pas organisées selon un modèle connu (comme une image ou un extrait de texte).

Raisonnement/traitement des informations et prise de décisions. Le module de raisonnement/traitement des informations est l'élément central d'un système d'IA. Il utilise les données provenant des capteurs et propose une action à prendre, en fonction de l'objectif à atteindre. Cela signifie que les données collectées par les capteurs doivent être transformées en informations compréhensibles par le module de raisonnement/traitement des informations. Pour développer notre exemple d'un système de nettoyage fondé sur l'IA, l'appareil photo fournira une photo du sol au module de raisonnement/traitement des informations, qui devra ensuite décider de nettoyer ou non le sol (en d'autres termes, il devra déterminer l'action la plus adaptée pour parvenir à l'objectif assigné). Bien qu'il puisse nous sembler simple, en tant qu'êtres humains, de décider de nettoyer ou non le sol à partir d'une photo, ce l'est beaucoup moins pour une machine, car une photo n'est qu'une séquence de 0 et de 1. Le module de raisonnement/traitement des informations doit par conséquent :

1. Interpréter la photo pour décider si le sol est propre ou non. De manière générale, cela signifie être capable de transformer des données en informations et de modéliser ces informations de manière succincte, en préservant toutefois tous les éléments de données pertinents (dans ce cas-ci, la question de savoir si le sol est propre ou non).
2. Appliquer un raisonnement à ces connaissances ou traiter ces informations pour produire un modèle numérique (c'est-à-dire une formule mathématique) afin de décider de la meilleure action à prendre. Dans cet exemple, si

l'information dérivée de la photo est que le sol est sale, l'action la plus adaptée consiste à activer le nettoyage, tandis que dans le cas contraire, l'action la plus adaptée serait de rester en veille.

Il est à souligner que le terme «décision» doit être considéré au sens large, comme tout acte consistant à choisir l'action à prendre, et ne signifie pas nécessairement que les systèmes d'IA sont complètement autonomes. Une décision peut également être le choix d'une recommandation à présenter à un être humain, qui sera le décideur final.

Actionnement. Lorsque l'action a été décidée, le système d'IA est prêt à la réaliser grâce aux actionneurs dont il dispose. Dans l'illustration ci-dessus, les actionneurs sont représentés par des bras articulés, mais il ne doit pas nécessairement s'agir d'éléments physiques. Les actionneurs peuvent également être des logiciels. Dans notre exemple concernant le nettoyage, le système d'IA pourrait émettre un signal qui active un aspirateur si l'action consiste à nettoyer le sol. Pour prendre un autre exemple, un système conversationnel (communément appelé «chatbot») agit en générant des textes pour répondre aux messages d'un utilisateur.

L'action réalisée sera peut-être de nature à modifier l'environnement, de telle sorte que, la prochaine fois que le système utilisera à nouveau ses capteurs, les informations qui lui parviendront seront peut-être différentes.

Les systèmes d'IA rationnels ne choisissent pas toujours l'action qui est la meilleure par rapport à leur objectif et ne parviennent donc à atteindre qu'une *rationalité limitée*, en raison de la limitation de ressources telles que le temps ou la puissance de calcul.

Les *systèmes d'IA rationnels* sont une version très élémentaire des systèmes d'IA. Ils modifient l'environnement mais n'adaptent pas leur comportement au fil du temps pour mieux atteindre leur objectif. Un *système rationnel apprenant* est un système rationnel qui, après avoir exécuté une action, évalue le nouvel état de l'environnement (par perception) pour déterminer le niveau de succès de son action et adapte ensuite ses règles de raisonnement et ses méthodes de prise de décisions.

2. L'IA en tant que discipline scientifique

La description d'un système d'IA présentée ci-dessus, qui est à la fois très simple et très abstraite, se fonde sur trois capacités principales: la perception, le raisonnement/la prise de décisions, et l'actionnement. Cette description est toutefois suffisante pour présenter et faire comprendre la plupart des techniques et sous-disciplines de l'IA actuellement utilisées dans l'élaboration de systèmes d'IA, car elles font toutes références aux différentes capacités des systèmes. De manière générale, toutes ces techniques peuvent être regroupées en deux groupes principaux qui renvoient à la capacité de *raisonnement* et d'*apprentissage*. La robotique est une autre discipline tout à fait pertinente.

Raisonnement et prise de décisions. Ce groupe de techniques comprend la représentation des connaissances ainsi que le raisonnement, la planification, la programmation, la recherche et l'optimisation. Ces techniques permettent d'appliquer un raisonnement aux données provenant des capteurs. Pour y parvenir, il convient de transformer les données en connaissances; un domaine de l'IA porte par conséquent sur la meilleure manière de modéliser ces connaissances (*représentation des connaissances*). Une fois les connaissances modélisées, l'étape suivante consiste à leur appliquer un raisonnement (*raisonnement sur les connaissances*). Cette étape passe par des déductions sur la base de règles symboliques, des activités de *planification* et de *programmation*, de la *recherche* parmi un vaste ensemble de solutions, et de l'*optimisation* parmi l'éventail de solutions possibles pour résoudre un problème. La dernière étape consiste à décider quelle action il convient de prendre. La partie relative au raisonnement/à la prise de décisions d'un système d'IA est en général très complexe et requiert une combinaison de plusieurs des techniques susmentionnées.

Apprentissage. Ce groupe de techniques comprend l'apprentissage automatique, les réseaux neuronaux, l'apprentissage profond, les arbres décisionnels et de nombreuses autres techniques d'apprentissage. Ces techniques permettent à un système d'IA d'apprendre à résoudre des problèmes qu'il n'est pas possible de spécifier précisément ou dont les méthodes de résolution ne peuvent être décrites par des règles symboliques de

raisonnement. Les exemples de tels problèmes sont ceux qui concernent les capacités de perception telles que la *compréhension de la parole et du langage*, ainsi que la *vision artificielle* ou la *prédiction des comportements*. En apparence, ces problèmes sont simples, car ils le sont en effet généralement pour les êtres humains. Ils ne sont toutefois pas si simples à résoudre pour les systèmes d'IA, ces derniers n'étant pas capables de fonder leur raisonnement sur le « bon sens » (du moins pas encore), et ils sont particulièrement compliqués lorsque le système doit interpréter des données non structurées. C'est ici que les techniques qui suivent l'approche de l'*apprentissage automatique* sont utiles. Toutefois, les techniques d'apprentissage automatique peuvent être appliquées à de nombreuses autres tâches que la seule perception. Les techniques d'apprentissage automatique produisent un modèle numérique (c'est-à-dire une formule mathématique) servant à calculer la décision à partir des données.

L'apprentissage automatique existe sous différentes formes. Les approches les plus répandues sont l'*apprentissage supervisé*, l'*apprentissage non supervisé* et l'*apprentissage par renforcement*.

Dans l'apprentissage automatique supervisé, au lieu de dicter des règles de conduite au système, nous lui fournissons des exemples de comportement d'entrée/de sortie, dans l'espoir qu'il soit capable de procéder à des généralisations à partir de ces exemples (décrivant souvent des situations passées) et de bien se comporter également dans des situations non présentées dans les exemples (qui pourraient être rencontrées dans le futur). Dans l'exemple qui nous occupe, nous fournirions au système de nombreux exemples de photos d'un sol accompagnés de l'interprétation correspondante (à savoir, si le sol est propre ou non sur cette photo). En fournissant suffisamment d'exemples, variés et suffisamment représentatifs de la plupart des situations, le système, au moyen de son algorithme d'apprentissage automatique, sera capable de procéder à des généralisations pour savoir de quelle manière interpréter correctement des photos de sols qu'il n'a jamais vues auparavant. Certaines approches d'apprentissage automatique adoptent des algorithmes fondés sur le concept des *réseaux neuronaux*, qui s'inspirent plus ou moins du cerveau humain, en ce sens qu'ils sont composés d'un réseau de petites unités de traitement (de manière analogue à nos neurones) reliées entre elles par de nombreuses connexions pondérées. Un réseau neuronal reçoit en entrée les données provenant des capteurs (dans notre exemple, la photo du sol) et a comme sortie l'interprétation de la photo (dans notre exemple, si le sol est propre ou non). Au cours de l'analyse des exemples (la phase d'*entraînement* du réseau), les pondérations des connexions sont ajustées pour correspondre autant que possible aux exemples disponibles (à savoir, réduire le plus possible l'erreur entre le résultat attendu et le résultat calculé par le réseau). À la suite de la phase d'entraînement, une phase d'essai du comportement du réseau neuronal par rapport à des exemples qu'il n'a jamais vus auparavant vérifie que la tâche a été correctement apprise.

Il convient de noter que cette approche (comme toute technique d'apprentissage automatique) présente toujours un certain pourcentage d'erreur, même s'il est généralement faible. La *précision*, une mesure du pourcentage de réponses correctes, est donc une notion essentielle.

Il existe plusieurs types de réseaux neuronaux et d'approches d'apprentissage automatique. Parmi celles qui donnent actuellement les meilleurs résultats figure l'*apprentissage profond*. Cette approche fait référence au fait que le réseau neuronal est composé de plusieurs couches entre l'entrée et la sortie, qui permettent un apprentissage de la relation générale entrée-sortie par étapes successives. Cela rend l'approche générale plus précise et moins dépendante des instructions données par l'être humain.

Les réseaux neuronaux ne sont qu'un outil d'apprentissage automatique; il en existe de nombreux autres, dotés de différentes propriétés: les forêts d'arbres décisionnels et le boosting d'arbres de décision, les méthodes de regroupement, la factorisation matricielle, etc.

L'*apprentissage par renforcement* est une autre approche d'apprentissage automatique utile. Dans cette approche, nous laissons le système d'IA prendre ses décisions et nous lui envoyons à chaque décision un signal de récompense lui indiquant si la décision était bonne ou mauvaise. Ce système a pour but, au fil du temps, de maximiser les récompenses positives obtenues. Cette approche est par exemple utilisée dans les systèmes de recommandation (tels que de nombreux systèmes de recommandation en ligne qui présentent aux utilisateurs des suggestions d'achat) ou encore dans le marketing.

Les approches d'apprentissage automatique sont non seulement utiles dans les tâches de perception, telles que la vision et la compréhension de textes, mais également dans toutes les tâches difficiles à définir et ne pouvant pas être décrites de manière exhaustive par des règles symboliques de comportement.

Une distinction est à observer entre les approches d'apprentissage automatique pour apprendre une nouvelle tâche qui ne peut être précisément décrite d'une façon symbolique, et les agents rationnels apprenants (mentionnés à la section précédente) qui adaptent leur comportement au fil du temps pour mieux atteindre l'objectif donné. Ces deux techniques peuvent se chevaucher ou coopérer, mais ne sont pas nécessairement identiques.

Robotique. La robotique peut être définie comme «l'IA en action dans le monde réel» (également appelée *IA embarquée*). Un robot est une machine physique qui doit s'accommoder aux dynamiques, aux incertitudes et à la complexité du monde réel. Des capacités de perception, de raisonnement, d'action, d'apprentissage ainsi que d'interaction avec d'autres systèmes sont en général intégrées à l'architecture de commande du système robotique. En plus de l'IA, d'autres disciplines interviennent dans la conception et le fonctionnement des robots, telles que l'ingénierie mécanique et la théorie du contrôle. Parmi les exemples de robots figurent les robots manipulateurs, les véhicules autonomes (voitures, drones, taxis volants, etc.), les robots humanoïdes, les aspirateurs-robots, etc.

La figure 2 représente la plupart des sous-disciplines de l'IA mentionnées plus haut, ainsi que leurs relations. L'IA est toutefois bien plus complexe que le montre cette illustration, car elle comprend de nombreuses autres sous-disciplines et techniques. Qui plus est, comme indiqué plus haut, la robotique s'appuie également sur des techniques qui sortent du champ de l'IA. Nous estimons cependant que ces informations sont suffisantes pour apporter un éclairage utile au partage, à la sensibilisation et aux discussions dans le domaine de l'IA, de l'éthique et des politiques en matière d'IA qui doivent avoir lieu au sein du groupe d'experts de haut niveau, fortement pluridisciplinaire et multipartite.

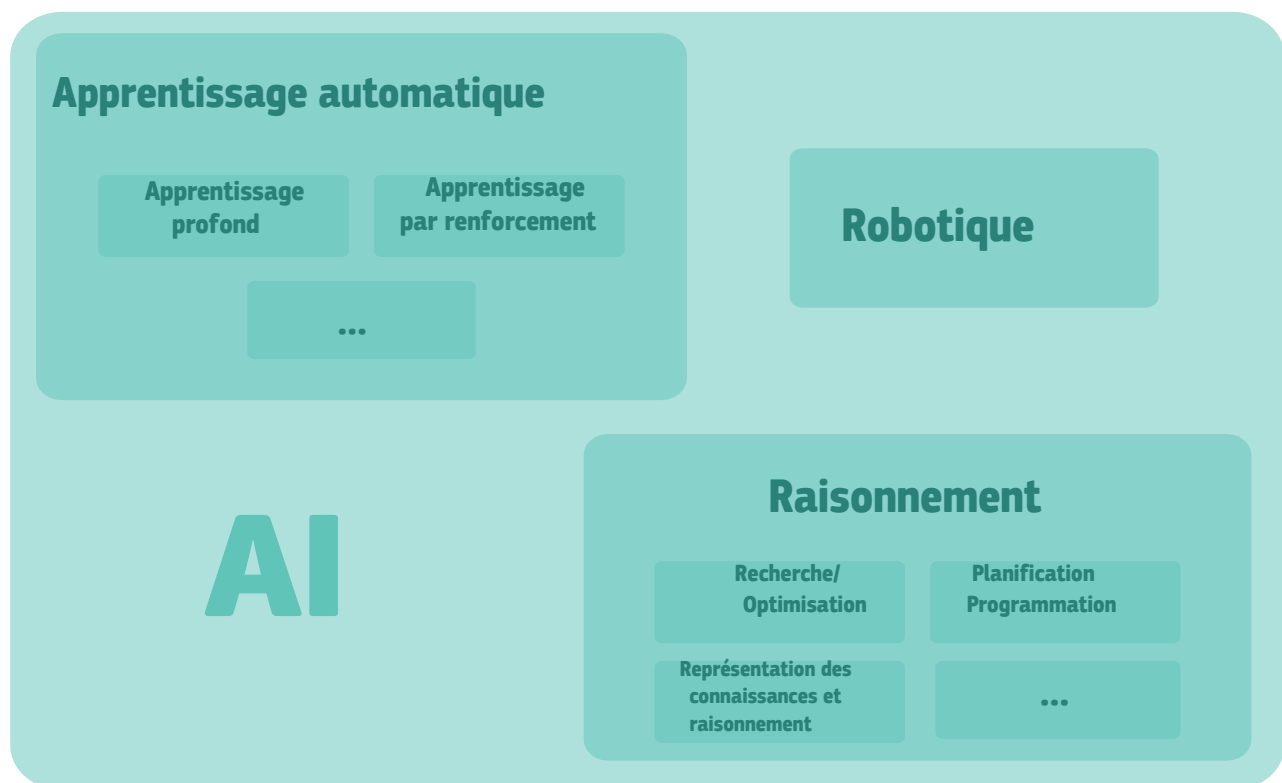


Figure 2: Aperçu simplifié des sous-disciplines de l'IA et de leurs relations.

Tant l'apprentissage automatique que le raisonnement automatique comprennent de nombreuses autres techniques, et la robotique comprend des techniques qui ne relèvent pas de l'IA. L'IA dans son ensemble relève de la discipline de l'informatique.

3. Autres notions et questions importantes dans le domaine de l'IA

IA étroite (ou faible) et générale (ou forte). Un système d'IA générale a vocation à réaliser la plupart des activités que les humains peuvent mener à bien. À l'inverse, les systèmes d'IA étroite peuvent réaliser une seule tâche ou quelques tâches spécifiques. Les systèmes d'IA actuellement déployés sont des exemples d'IA étroite. À l'époque des débuts de l'IA, les chercheurs utilisaient une terminologie différente (IA faible et forte). De nombreux défis éthiques, scientifiques et technologiques restent à relever pour développer les capacités qui seraient nécessaires pour réaliser une IA générale, par exemple le raisonnement de sens commun, la conscience de soi et la capacité d'un système à définir ses propres objectifs.

Problèmes liés aux données et biais. Étant donné que de nombreux systèmes d'IA, tels que ceux qui comprennent des éléments d'apprentissage automatique supervisé, ont besoin d'énormes quantités de données pour fonctionner correctement, il est important de comprendre la manière dont les données influent sur le comportement du système d'IA. Par exemple, si les données d'entraînement sont biaisées, c'est-à-dire qu'elles ne sont pas suffisamment équilibrées ou inclusives, le système d'IA entraîné avec ce type de données ne sera pas capable de procéder à une généralisation satisfaisante et prendra peut-être des décisions injustes susceptibles de favoriser certains groupes par rapport à d'autres. La communauté de l'IA travaille depuis peu à des méthodes visant à détecter et à atténuer les biais dans les ensembles de données d'entraînement ainsi que dans d'autres parties d'un système d'IA.

Boîte noire de l'IA et explicabilité. Certaines techniques d'apprentissage automatique, tout en donnant de très bons résultats du point de vue de la précision, sont très opaques pour ce qui est de comprendre la manière dont elles prennent leurs décisions. La notion de *boîte noire de l'IA* fait référence à de tels scénarios, dans lesquels il n'est pas possible de déterminer la raison de certaines décisions. À l'inverse, l'explicabilité est une propriété des systèmes d'IA qui sont, eux, capables de fournir une forme d'explication de leurs actions.

IA dirigée par le but. Les systèmes d'IA actuels sont dirigés par le but, ce qui signifie qu'un objectif à atteindre leur est précisé par un être humain et qu'ils utilisent certaines techniques pour parvenir à cet objectif. Ils ne définissent pas leurs propres objectifs. Toutefois, certains systèmes d'IA (tels que ceux qui se fondent sur certaines techniques d'apprentissage automatique) peuvent avoir plus de liberté pour décider quelle voie emprunter pour réaliser l'objectif donné.

4. Définition actualisée de l'IA

Nous proposons d'utiliser la définition actualisée suivante de l'IA:

«Les systèmes d'intelligence artificielle (IA) sont des systèmes logiciels (et éventuellement matériels) conçus par des êtres humains³ et qui, ayant reçu un objectif complexe, agissent dans le monde réel ou numérique en percevant leur environnement par l'acquisition de données, en interprétant les données structurées ou non structurées collectées, en appliquant un raisonnement aux connaissances, ou en traitant les informations, dérivées de ces données et en décidant de la/des meilleure(s) action(s) à prendre pour atteindre l'objectif donné. Les systèmes d'IA peuvent soit utiliser des règles symboliques, soit apprendre un modèle numérique. Ils peuvent également adapter leur comportement en analysant la manière dont l'environnement est affecté par leurs actions antérieures.

En tant que discipline scientifique, l'IA comprend plusieurs approches et techniques, telles que l'apprentissage automatique (dont l'apprentissage profond et l'apprentissage par renforcement sont des exemples spécifiques), le raisonnement automatique (qui comprend la planification, la programmation, la représentation des connaissances et le raisonnement, la recherche et l'optimisation) et la robotique (qui comprend le contrôle, la perception, des capteurs et des actionneurs, ainsi que l'intégration de toutes les autres techniques dans des systèmes cyberphysiques).».

³ Les systèmes d'IA sont conçus directement par des êtres humains mais ces derniers peuvent aussi recourir à des techniques relevant de l'IA pour optimiser la conception de ces systèmes..

et de renvoyer au présent document en tant que source d'informations complémentaires pour soutenir cette définition.

**Le présent document a été élaboré par les membres du groupe d'experts de haut niveau sur
l'IA**

énumérés ci-dessous par ordre alphabétique

Pekka Ala-Pietilä, président du GEHN IA AI Finland, Huhtamaki, Sanoma	Pierre Lucas Orgalim – Europe's technology industries
Wilhelm Bauer Fraunhofer	Ieva Martinkenaite Telenor
Urs Bergmann Zalando	Thomas Metzinger JGU Mainz & European University Association
Mária Bielíková Slovak University of Technology in Bratislava	Catelijne Muller ALLAI Netherlands & EESC
Cecilia Bonefeld-Dahl DigitalEurope	Markus Noga SAP
Yann Bonnet ANSSI	Barry O'Sullivan, vice-président du GEHN IA University College Cork
Loubna Bouarfa OKRA	Ursula Pacht BEUC
Stéphan Brunessaux Airbus	Nicolas Petit Université de Liège
Raja Chatila Initiative de l'IEEE Ethics of Intelligent/Autonomous Systems & Sorbonne Université	Christoph Peylo Bosch
Mark Coeckelbergh Université de Vienne	Iris Plöger BDI
Virginia Dignum Université d'Umeå	Stefano Quintarelli Garden Ventures
Luciano Floridi Université d'Oxford	Andrea Renda College of Europe Faculty & CEPS
Jean-François Gagné Element AI	Francesca Rossi* IBM
Chiara Giovannini ANEC	Cristina San José Fédération bancaire de l'Union européenne
Joanna Goodey Agence des droits fondamentaux	George Sharkov Digital SME Alliance
Sami Haddadin Munich School of Robotics and MI	Philipp Slusallek German Research Centre for AI (DFKI)
Gry Hasselbalch The thinkdotank DataEthics & Université de Copenhague	Françoise Soulié Fogelman Consultante en IA
Fredrik Heintz Université de Linköping	Saskia Steinacker Bayer
Fanny Hidvegi Access Now	Jaan Tallinn Ambient Sound Investment
Eric Hilgendorf Université de Würzburg	Thierry Tingaud STMicroelectronics
Klaus Höckner Hilfsgemeinschaft der Blinden und Sehschwachen	Jakob Uszkoreit Google
Mari-Noëlle Jégo-Laveissière Orange	Aimee Van Wynsberghe TU Delft
Leo Kärkkäinen Nokia Bell Labs	Thiébaut Weber ETUC
Sabine Theresia Köszegi TU Wien	Cecile Wendling AXA
Robert Kroplewski Avocat & conseiller du gouvernement polonais	Karen Yeung Université de Birmingham
Elisabeth Ling RELX	

*Francesca Rossi a assumé le rôle de rapporteur pour le présent document.