



**EXAMEN PROFESSIONNEL
POUR LE RECRUTEMENT DANS LE GRADE DE SECRETAIRE
D'ADMINISTRATION DE CLASSE SUPERIEURE DE LA CAISSE DES
DEPOTS ET CONSIGNATIONS**

SESSION 2021

Mardi 10 mars 2020

L'épreuve d'admissibilité consiste à traiter un cas pratique, destiné à mettre le candidat en situation de travail, à partir d'un dossier, qui est assorti de questions. Ce dossier, à caractère administratif, ne pouvant excéder trente pages, peut comporter des graphiques et des données chiffrées (durée : 3 heures ; coefficient 2).

Important :

Vous devez écrire uniquement sur la copie d'examen qui vous a été remise et sur les **Intercalaires qui vous seront distribués si besoin.**

Si vous utilisez des feuilles intercalaires vous devez inscrire la pagination en haut à droite de chaque feuille et reporter votre numéro de « code à barres ».



Sous peine de nullité, votre copie et vos intercalaires, ne doivent en aucun cas être signés ou comporter un signe distinctif permettant l'identification du candidat (signature, nom, paraphe, initiales, symbole, ...).

Vous devez écrire à l'encre bleue ou noire – pas d'autre couleur– sous peine de nullité car cela peut s'apparenter à un signe distinctif.

Il vous est demandé d'apporter un soin particulier à la présentation de votre copie.

Le dossier contient 30 pages

Sujet

Vous êtes secrétaire administratif au Ministère de l'enseignement supérieur, de la recherche et de l'innovation. Votre directeur doit intervenir prochainement, pour le Ministère, au cours d'un colloque portant sur les innovations technologiques et en particulier sur le thème de l'intelligence artificielle : « danger ou progrès ».

Il vous demande de lui rédiger une note synthétique qui présentera :

- **Quelles sont les problématiques liées à l'intelligence artificielle et plus largement aux algorithmes ? Quels sont les enjeux, en particulier dans la recherche, liés à l'intelligence artificielle ?**
- **Les dangers potentiels de l'intelligence artificielle auxquels il faudrait réfléchir dès aujourd'hui.**
- **Les espoirs fondés dans l'intelligence artificielle, en mettant en perspective les recommandations préconisées sur l'intelligence artificielle et plus largement les algorithmes.**

Vous accompagnerez votre note d'un power point d'une page présentant brièvement, de façon thématique, les applications actuelles de l'intelligence artificielle en étayant avec quelques données chiffrées éclairantes pour votre directeur.

Pour mener à bien ce travail, vous disposez des documents ci-après.

Liste des documents

Document 1 : Site internet Ville de Paris « L'intelligence artificielle au service du trafic urbain » Posté le 9 Janvier 2020 (1 page)

Document 2 : Les échos « L'imagerie médicale à l'heure de l'intelligence artificielle » 17 décembre 2019 (2 pages)

Document 3 : Les Echos « Huggins Face, l'intelligence artificielle en open source » 18 décembre 2019 (2 pages)

Document 4 : Le Monde « L'intelligence artificielle, un nouveau secteur à investir » 14 décembre 2019 (2 pages)

Document 5 : Le Monde « Fausses informations : un programme d'IA jugé trop dangereux pour être rendu public » Posté le 19 Février 2019 (3 pages)

Document 6 : Le Parisien « Intelligence artificielle | Macron annonce un plan à 1,5 milliard d'euros » - Posté le 29 mars 2018 (1 page)

Document 7 : Synthèse du rapport Villani – Posté sur le site de AI for Humanity Mars 2018 (10 pages)

Document 8 : CNIL Extrait du rapport « Comment permettre à l'homme de garder la main - Les enjeux éthiques des algorithmes et de l'intelligence artificielle » 15 Décembre 2017 (4 pages)

Document 9 : Le Monde « Une intelligence artificielle peut-elle devenir présidente des Etats-Unis ? » Posté le 17 Février 2016 (2 pages)

Document 10 : France 24 « L'intelligence artificielle inquiète les grands cerveaux contemporains » - 2 Février 2015 (2 pages)

Document 11 : Encyclopédie Larousse « Définition de l'intelligence artificielle » (1 page)

Document 1

Wintics - L'intelligence artificielle au service du trafic urbain

Site internet Ville de Paris posté le 09/01/2020

Récompensée aux Grands Prix de l'Innovation 2019, cette technologie analyse les images captées par caméras grâce à l'intelligence artificielle. Les statistiques récupérées permettent de mieux comprendre le trafic en temps réel afin de le fluidifier. Créé par Quentin Barenne, 32 ans, Matthias Houllier, 31 ans et Levi Viana, 33 ans, Wintics ambitionne de rendre la ville moins congestionnée et moins polluée. Depuis décembre 2017, la start-up développe cette solution auprès des clients, entreprises, villes ou gestionnaires pour mieux connaître et analyser les flux urbains. Une innovation récompensée dans la catégorie "énergie et mobilité" des grands prix de l'innovation 2019 de la Ville de Paris.

Analyser les Itinéraires parcourus

Branchée sur les caméras de vidéosurveillance déjà installées, le logiciel « Smart Caméra », qui peut être hébergé dans des petits boîtiers de la taille d'une canette, enregistre les allers et venues des carrefours urbains. *« Le logiciel peut se brancher sur n'importe quelle caméra, qu'elle soit optique ou thermique, indique Quentin Barenne, cofondateur de la start-up. Il capte tous les mouvements des piétons, des cyclistes et des voitures. »*

Ces données, des images brutes, sont transformées en statistiques au sein même du boîtier grâce à des algorithmes d'intelligence artificielle. Le système de vision par ordinateur est ainsi capable de détecter les différents moyens de transports et de comprendre les itinéraires parcourus à partir des images (qui font apparaître des formes pour chaque usager et non des visages).

Les clients reçoivent alors uniquement les statistiques, en aucun cas les vidéos. Les chiffres récoltés renseignent sur les volumes et les directions des piétons et des véhicules sur les voies publiques, les gares, les parkings ou encore les centres commerciaux. Les applications des données recueillies sont ensuite multiples : elles dépendent de la commande initiale du client, des gestionnaires de voirie, collectivités, opérateurs de transport et équipementiers.

Favoriser les véhicules avec passagers

« L'enjeu est double : fluidifier le trafic et sécuriser les piétons et les cyclistes », explique Quentin Barenne. Avec ces données, les clients vont pouvoir adapter leurs services. « Par exemple, avec les caméras à l'entrée d'un parking, on va détecter le nombre de passagers à bord d'une voiture. Cela va permettre au gestionnaire d'appliquer des réductions aux véhicules avec plus de passagers », poursuit le cofondateur de Wintics.

Un tiers du trafic dans les grandes villes en heures de pointe est lié aux voitures en recherche d'une place

La solution peut aussi identifier en temps réel les places de stationnement libre dans une rue. Une avancée « significative » pour Wintics car *« un tiers du trafic dans les grandes villes en heures de pointe est lié aux voitures en recherche d'une place »*. Les données produites intéressent aussi les gestionnaires de transport public afin de les aider à comprendre comment leurs infrastructures et réseaux sont utilisés par les usagers. Alors souriez, vous êtes analysés !

Document 2

Les échos « L'imagerie médicale à l'heure de l'intelligence artificielle »

17/12/2019

Les premiers modules d'intelligence artificielle sont maintenant proposés par les constructeurs.

Ils aident à interpréter les images et choisir les traitements mais aussi à gérer les patients.

Il y a deux ans déjà, l'intelligence artificielle (IA) faisait irruption à la Une du Congrès de la Radiological Society of North America (RSNA), le grand rendez-vous mondial annuel de l'imagerie médicale, avec ses quelque 50.000 participants. Une peur millénariste avait alors saisi une profession redoutant sa mort prochaine, au profit des machines devenues plus intelligentes que les humains. Du fantôme à la réalité, avec les premières réalisations concrètes en cette fin 2019, de nombreux professionnels ont réalisé qu'ils pouvaient s'emparer de l'IA pour alléger leur quotidien.

Car les chiffres sont là. « En 2010, il fallait trois ans et demi pour que la quantité de données médicales soit multipliée par deux. En 2020, il ne faut plus que 72 jours », observe le directeur médical de GE Healthcare Europe, Mathias Goyen. Et l'imagerie y contribue largement. Elle est non seulement utilisée pour le diagnostic dans un nombre croissant de situations, mais aussi pour le suivi des traitements et même leur réalisation avec l'essor de la radiologie interventionnelle.

Face à cette explosion des données, « la ressource critique aujourd'hui, ce sont les professionnels de santé », observe David Corcos, qui dirige Philips en France. « Il en manquerait 7 millions dans le monde actuellement, chiffre qui devrait passer à 14 millions en 2035 », selon Mathias Goyen. « Et pour ceux qui exercent, le burn-out guette », s'inquiète le patron du service d'imagerie de l'hôpital Foch à Paris, Philippe Grenier.

Interprétation d'images

Dans ce contexte, l'intelligence artificielle n'est pas la panacée, mais elle constitue une aide dont il sera de plus en plus difficile de se passer. Tous les constructeurs d'équipements d'imagerie l'ont compris. Ils continuent certes à innover dans le « hardware ». Canon revendique ainsi les plus petits détecteurs pour son nouveau scanner, Philips a mis au point un scanner à comptage photonique et une IRM sans hélium, ou encore Siemens a développé son IRM 7 Teslas, en routine à Poitiers depuis novembre. Mais une part croissante de la R&D est désormais tournée vers le « software ».

« Siemens dispose à Princeton d'une grosse infrastructure d'intelligence artificielle, employant quelque 600 ingénieurs et scientifiques. Il a été le premier, à obtenir la validation de la FDA de son module d'interprétation d'images scanner IA-Rad Companion, pour son application au thorax », explique Christoph Zindel, membre du comité exécutif de Siemens Healthineers chargé de l'intelligence artificielle. Ce module, qui le dispute à l'AICE de Canon, attire l'attention du médecin sur les

anomalies éventuellement décelées sur les différents organes et prévient la baisse de vigilance du radiologue résultant de la fatigue.

Choix du traitement

La validation clinique des outils d'intelligence artificielle suppose des partenariats avec les hôpitaux. Philips a signé en France pour deux ans avec le CHU de Lille, pour l'évaluation d'illumeo, un module qui rassemble, pour un patient, la liste de ces pathologies, ses anciens examens d'imagerie, ses résultats d'analyses biologiques et ses prescriptions, issus d'autres bases de données que les images. Avec son Navify Tumor Board, développé en collaboration avec Roche, GE Healthcare propose un produit dans le même esprit pour les équipes de cancérologie, afin de faciliter le choix du traitement. Idem chez Siemens, où le IA-Pathway companion fournit un tableau de bord de la tumeur, avec comme première application le cancer de la prostate.

Toutes ces utilisations de l'IA visent à raccourcir la durée des examens et à optimiser les flux de patients, tout en assurant une qualité homogène des soins. Les grands constructeurs ne sont pas les seuls à développer de tels modules. A côté d'eux, une myriade de petites sociétés y travaillent aussi. C'est ce qui a incité Philips à créer IntelliSpace, une plate-forme ouverte à travers laquelle, elles acquièrent une visibilité auprès des utilisateurs.

Supervision à distance

Enfin, comme dans d'autres domaines, l'intelligence artificielle permet de gérer et d'assurer la maintenance à distance des parcs de scanner, d'IRM ou autres... mais aussi de surveiller des patients. Face au turn-over important du personnel médical, la chaîne d'hôpitaux Advocate Aurora Health a mis en place, aux Etats-Unis, un système de supervision à distance.

Document 3

Les Echos « Huggins Face, l'intelligence artificielle en open source »

18/12/2019

Hugging Face, l'intelligence artificielle en open source, lève 15 millions

La start-up new-yorkaise créée par deux Français a modifié son modèle il y a un an. Elle développe aujourd'hui une technologie open source de traitement du langage naturel, utilisée par un millier d'entreprises. Pour accélérer ses progrès, elle vient de lever 15 millions de dollars.

L'intelligence artificielle n'est pas l'affaire des seuls Google, Amazon ou Facebook. Les Français de Hugging Face en sont persuadés. La jeune pousse, dont les fondateurs et un tiers des équipes sont installés à New York, veut fédérer une communauté open source autour du traitement du langage naturel, une des composantes de l'intelligence artificielle.

A l'origine, Hugging Face avait créé un ami virtuel, un chatbot avec qui l'on pouvait interagir. « Il fallait comprendre le sens, le sentiment du message, répondre, générer du texte. Il s'agissait de tâches complètes qui nous ont amenés à travailler sur une des meilleures technologies de traitement du langage naturel », affirme Clément Delangue, le cofondateur de Hugging Face. C'est ainsi qu'en octobre 2018 la société a lancé sa technologie en open source, « ce qui a permis une adoption exceptionnelle, très rapide », explique l'entrepreneur.

Aujourd'hui, un millier d'entreprises utilise ainsi la technologie de Hugging Face. C'est le cas de Microsoft, sur son moteur de recherche Bing. Quand un utilisateur y pose une question en langage naturel (par exemple « qui était le 34^e président des Etats-Unis ? »), Bing utilise la technologie de la jeune pousse pour extraire l'information parmi les pages les plus pertinentes. Des médias s'en servent aussi pour détecter les informations essentielles dans la communication financière d'une entreprise, par exemple, pour commencer une dépêche. D'autres entreprises, enfin, trient les informations contenues dans les e-mails reçus pour aiguiller le service clients.

Tripler les effectifs

« Nous sommes persuadés que l'intelligence artificielle va se développer grâce à l'open source, et non seulement avec quelques géants travaillant dans leur coin », assure Clément Delangue. Pour accélérer le développement de sa technologie, Hugging Face vient de lever 15 millions de dollars, portant le total des fonds levés à près de 20 millions.

Le fonds américain Lux Capital, spécialisé dans les projets scientifiques, fait son entrée au capital en prenant une participation minoritaire. C'est lui qui avait investi, entre autres, dans CTRL-Labs, racheté récemment par Facebook, ou dans la biotech Auris Health, acquise par Johnson & Johnson pour 3,4 milliards de dollars. Il rejoint les investisseurs historiques, comme Betaworks et quelques business angels, comme le directeur scientifique de Salesforce, Richard Socher, l'entrepreneur Thibaud Elzière, ou le basketteur des Brooklyn Nets Kevin Durant, qui a remis un ticket dans ce tour de table.

Hugging Face veut tripler ses effectifs et passer à une trentaine de personnes : une dizaine à Paris et une vingtaine à New York. La société veut recruter des ingénieurs et des scientifiques, des profils de plus en plus recherchés, en plus des quelque 200 contributeurs que sa communauté compte à travers le monde.

Si la société ne génère pas de revenus pour l'instant, cela n'inquiète pas les investisseurs. « Ce n'est pas une priorité. Si l'on crée la technologie la plus utilisée, on crée déjà beaucoup de valeur », estime Clément Delangue, qui pense que si la science continue de progresser à la même vitesse, les interactions générées seront du niveau d'un humain d'ici à trois ans.

Document 4

Le Monde « L'intelligence artificielle, un nouveau secteur à investir »

14 décembre 2019

Issue de la révolution numérique, l'intelligence artificielle (IA) est en plein boom. Permettant aux machines et aux logiciels de penser comme un être humain, cette jeune technologie commence à bouleverser les métiers des nombreuses entreprises et pourrait leur ouvrir de nouvelles perspectives de développement.

Voici pourquoi les gestionnaires d'actifs s'intéressent à cette thématique innovante qui devrait rester dans l'ère du temps pendant un bon moment. *« Cette tendance de fond n'en est qu'à ses débuts. C'est une révolution silencieuse qui touche tous les secteurs »*, affirme Brice Prunas, gérant du fonds Intelligence artificielle d'Oddo BHF. Pour un épargnant attiré par ce thème, il est possible de faire une incursion boursière dans le monde de demain. Les fonds estampillés « IA » – Amundi, Oddo BHF, La Financière de l'Echiquier (LFDE), Allianz Global Investors, Candriam... – se comptent encore sur les doigts de la main. Cogérant du fonds Echiquier Artificial Intelligence lancé en 2018, Rolando Grandi donne sa définition de l'IA : elle repose sur *« quatre piliers que sont le big data, le cloud, les applications et les réseaux neuronaux artificiels »*. Ces derniers sont en quelque sorte des cerveaux artificiels rendant les machines de plus en plus intelligentes grâce à l'apprentissage (*« machine learning », « deep learning »...*).

Le fonds d'Oddo BHF compose un portefeuille en ciblant les « producteurs » d'intelligence artificielle, à savoir les sociétés qui développent ces nouveaux outils. *« Cela va des fabricants de puces électroniques à celui des logiciels et d'algorithmes jusqu'à la société produisant les infrastructures comme les semi-conducteurs pour faire passer la 5G par exemple »*, souligne Brice Prunas.

Amérique du Nord et Asie

De son côté, Candriam intègre dans son univers d'investissement uniquement les entreprises utilisatrices de cette technologie. *« C'est encore difficile, car il n'est pas évident d'évaluer la contribution de l'IA aux revenus d'une entreprise »*, signale Johan van Der Biest, gestionnaire de fonds Candriam. Pictet AM a pris le parti d'axer son fonds sur le seul volet « big data ».

« Grace à cette technologie disruptive développée par des jeunes sociétés innovantes, le potentiel de croissance des résultats à moyen terme est important. On est même au début de cette phase de création de valeur », affirme Brice Prunas.

Pour faire leur marché, la plupart des gestionnaires procèdent à une sélection des titres dans tous les secteurs d'activité (santé, industrie, consommation, technologies de l'information...) et dans tous les pays du globe. Chez Oddo BHF, on indique réaliser un premier filtre de ces valeurs mondiales à l'aide de l'intelligence artificielle. *« Cet outil aide à détecter les entreprises qui entrent dans nos critères mais sont "hors des radars" des gestionnaires. Notre stratégie consiste de capter de la création de valeur indépendamment de la taille de la société cotée »*, explique Brice Prunas.

Certains gestionnaires optent pour les petites et moyennes capitalisations de jeunes sociétés, d'autres misent sur les sociétés plus importantes parfois même sur certains fonds combinent les deux. Globalement, les zones géographiques les plus représentées dans ces portefeuilles de trente à cinquante « lignes » sont majoritairement l'Amérique du Nord, suivi de loin par l'Asie. « *C'est normal ce sont là que les pays sont leaders dans le secteur où les entreprises* », relève Johan Van Der Biest, gestionnaire du fonds chez Candriam.

Quelques ETF (« Exchange Traded Funds » ou « fonds cotés en Bourse », appelés « trackers » par les spécialistes) – DWS, Amundi – dédiés à l'intelligence artificielle commencent à fleurir sur le marché. « *Pour couvrir cette thématique de la rupture technologique, nous jouons la diversification des titres et des secteurs* », précise Olivier Souliac, responsable des stratégies indicielles chez Xtrackers (DWS).

Les freins à l'expansion de ce jeune secteur ? Des entreprises souvent récentes donc encore peu rentables ou davantage de contraintes juridiques concernant la gestion et le contrôle des données qui pourraient alourdir les coûts de développement.

Cinq fonds spécialisés en intelligence artificielle

Nom du fonds (société de gestion)	Montant de la part en euro	Performance en %, depuis janv. 2019
Allianz Global Artificial Intelligence AT H2R (Allianz)	128,15	+ 19,48
Candriam Equities L Robotics & Innovative Technology Class C USD Cap (Candriam)	199,5	+ 29,65
Echiquier Artificial Intelligence A EUR (Financière de l'Echiquier)	110,59	+ 28,76
Amundi Stoxx Global Artificial Intelligence EUR COAI (Amundi)	54,37	+ 25,94
Oddo BHF Artificial Intelligence CR-EUR (Oddo BHF)	118,64	N.C.

SOURCE : MORNINGSTAR, NOVEMBRE 2019

Document n°5

Le Monde Fausses informations : un programme d'IA jugé trop dangereux pour être rendu public

19 février 2019

OpenAI a-t-il créé un programme d'intelligence artificielle (IA) trop dangereux pour être rendu public ? C'est en tout cas ce que prétend un laboratoire américain à but non lucratif, dans un article publié, jeudi 14 février, sur son site Internet. Et copieusement commenté ces derniers jours par la communauté des chercheurs en IA.

Le programme en question s'appelle GPT-2 et sert à générer automatiquement du texte. Il a été « entraîné » à partir de 40 gigabits de textes, issus de huit millions de pages Web. « Avec un objectif simple, explique OpenAI : *prédire le mot suivant, en se basant sur tous les mots précédents d'un texte.* » Il suffit donc de lui donner quelques mots, comme une phrase, pour qu'il rédige une suite. Ce programme « *s'adapte au style et au contenu* » des mots qu'on lui donne, ce qui permet, selon OpenAI, de « *générer une suite réaliste et cohérente sur le sujet choisi* » par l'utilisateur. Un roman de fantasy, une recette, un article de presse ou un pamphlet politique, par exemple.

Le résultat est impressionnant, assure l'organisation, qui estime que sa technologie représente ce qui se fait de mieux aujourd'hui dans le domaine. Au point de s'en inquiéter. Pour elle, les textes produits par GPT-2 sont tellement réalistes que cette technologie, placée entre de mauvaises mains, pourrait se révéler dangereuse.

L'organisation pointe notamment le risque que ce programme serve à rédiger à la chaîne des articles présentant de fausses informations ou des publications trompeuses sur les réseaux sociaux. Il pourrait aussi, craignent ces chercheurs, être utilisé pour usurper l'identité de quelqu'un en ligne. Pour OpenAI, ces avancées, « *combinées à celles de l'imagerie, du son et de la vidéo de synthèse* », facilitent grandement la création, à moindre coût, « *de faux contenus et de campagnes de désinformation* ». Elle a donc décidé de ne pas rendre son programme public, contrairement à son habitude.

De belles réussites, mais aussi des ratés

Que vaut vraiment cette technologie ? OpenAI a permis à quelques journalistes de la tester. « *Une IA écrit exactement comme moi. Préparez-vous à l'apocalypse des robots* », s'est émue une chroniqueuse du *Guardian*. Elle explique d'ailleurs qu'un article généré par cette technologie a été imprimé dans les pages du *Guardian*, pour montrer sa performance, mais la rédaction a fait le choix de ne pas le publier en ligne, « *parce qu'on ne trouvait aucun moyen d'empêcher qu'il ne soit considéré comme un véritable article, sorti de son contexte* ».

Le magazine *Wired*, spécialisé dans les nouvelles technologies, a lui aussi pu tester ce programme, aux résultats « *extrêmement réalistes* », selon lui. Il publie quelques exemples (en anglais), notamment un texte ayant pour sujet imposé Hillary Clinton, candidate malheureuse à la dernière élection présidentielle américaine, et le milliardaire américain George Soros. Le texte généré par GPT-2 les accuse de collusion et évoque un enregistrement audio dans lequel les deux personnalités complotent pour faire passer de fausses informations. Si le texte est imprécis et manque de cohérence, il semble les accuser avec beaucoup d'arguments et de faits : bien qu'ils soient entièrement faux, l'ensemble est assez bien réalisé pour qu'il puisse paraître plausible.

OpenAI émet toutefois quelques nuances aux réussites de sa technologie. L'organisation évoque par exemple de nombreux essais ratés, avec des textes trop répétitifs ou absurdes.

« Quand on lui donne des sujets qui sont énormément représentés dans les données (le Brexit, Miley Cyrus, le Seigneur des anneaux, etc.), il est capable de générer un texte correct dans 50 % des cas. Mais l'inverse est aussi vrai : sur des contenus très techniques ou ésotériques, les résultats sont mauvais. »

OpenAI note aussi la présence de passages absurdes (comme lorsque le programme parle de feu sous l'eau) ou de changements de sujets abrupts. Ce qui s'explique par le fait que cette technologie, bien qu'elle réussisse à imiter l'écrit humain, est incapable d'en comprendre le sens.

« Course au clic »

Le problème est qu'en refusant de rendre sa technologie publique OpenAI rend impossible, pour la communauté scientifique, d'évaluer l'efficacité réelle du système. Ce qui lui a valu de nombreuses critiques. « *Où peut-on trouver des preuves que votre système est bien capable de faire ça ?* », s'agace par exemple Anima Anandkumar, responsable de la recherche en apprentissage automatique chez Nvidia.

« Quels chercheurs indépendants ont analysé votre système ? Aucun. (...) Si vous pensez que votre système est si efficace, ouvrez-le aux chercheurs. Pas à des journalistes qui font la course au clic. »

Plusieurs ont dénoncé ce qu'ils considèrent être une opération de communication. Zachary Chase Lipton, enseignant à l'université Carnegie Mellon, estime que les résultats d'OpenAI ne méritent pas un tel emballement. « *Ils sont intéressants mais pas surprenants. (...) C'est la sorte d'avancée qu'on s'attendrait à voir dans un mois ou deux, de la part de n'importe lequel des dizaines de labos travaillant sur le langage* », écrit-il sur son site.

Il faut dire qu'OpenAI est un habitué des coups médiatiques. C'est même dans son ADN : l'organisation a été fondée en 2015 par Elon Musk, coutumier des sorties catastrophistes sur l'IA, dans le but de répondre à ces dangers en développant elle-même des programmes d'IA *« plus à même de bénéficier à l'humanité »*. Elon Musk s'est depuis mis en retrait, ce qu'il a rappelé samedi après la publication de l'article d'OpenAI.

« Nous ne sommes pas sûrs que ce soit la bonne décision »

Si l'organisation est critiquée, c'est aussi parce qu'elle avait promis de rendre le résultat de ses recherches accessible à tous, expliquant à l'époque qu'elle voulait *« créer de la valeur pour tout le monde, plutôt que pour des actionnaires »* – arguant que ces technologies ne devaient pas rester entre les mains de quelques entités riches et puissantes.

Mais, visiblement, elle juge le cas de GPT-2 différent. Pour OpenAI, *« la communauté de l'IA va devoir s'attaquer à la question de la publication »*, en mettant en place des standards. L'organisation considère que les gouvernements devraient aussi s'y intéresser, pour *« étudier plus systématiquement l'impact social »* que pourrait avoir la diffusion de certaines technologies.

« Nous ne sommes pas sûrs aujourd'hui que ce soit la bonne décision », tient toutefois à tempérer OpenAI, mais *« nous estimons que cette expérimentation servira d'exemple pour des débats plus nuancés »* sur la question de la mise à disposition ou non de telles technologies.

Document 6

Le Parisien « Intelligence artificielle ! Macron annonce un plan à 1,5 milliard d'euros » -

29 mars 2018

Un « plan Intelligence artificielle » va être déployé pour catapulter la France au meilleur niveau mondial en intelligence artificielle.

Il veut faire de la France l'un des leaders en matière d'intelligence artificielle (IA). Après la remise du rapport du mathématicien et député (LREM) Cédric Villani mercredi, Emmanuel Macron a martelé son ambition lors d'un discours donné jeudi après-midi au Collège de France. « Nous avons les talents, nous avons tout pour relever le défi de l'IA », s'est-il félicité.

Pour y parvenir, un plan Intelligence artificielle va être déployé jusqu'en 2022. L'objectif : faire émerger « un réseau d'instituts dédiés localisés dans quatre ou cinq endroits en France », « constituer un hub de recherche au meilleur niveau mondial en IA », favoriser les allers-retours des chercheurs entre le public et le privé...

« L'ensemble du plan Intelligence Artificielle (IA) représentera sur le quinquennat un effort spécifique et inédit de près de 1,5 Md€ de crédits publics », a précisé le chef de l'Etat. Avec une enveloppe de près de 400 M€ dévolue à des « appels à projet et de défis d'innovation de rupture ».

Les salaires des chercheurs ne seront pas doublés

Où l'Etat ira-t-il ponctionner cet argent ? Principalement dans le Fonds pour l'innovation et l'industrie de 10 Mds € mis en place depuis le début d'année et qui doit permettre d'injecter chaque année environ 250M€ dans des projets d'envergure.

L'Etat devrait aussi ponctionner dans le Fond de transformation de l'action publique et par redéploiement budgétaire, assure l'Elysée sans en préciser les contours précis. Emmanuel Macron a par ailleurs exclu de doubler les salaires des chercheurs, comme le souhaitait pourtant Cédric Villani.

D'autre part, le cadre législatif va évoluer pour permettre la mise en application de l'IA. Un exemple : dès l'an prochain, il sera possible d'expérimenter sur les routes les véhicules autonomes les plus perfectionnés. Et d'ici à 2022, un « cadre de régulation » sera instauré pour permettre leur circulation.

Preuve que la France séduit, plusieurs géants internationaux ont déjà annoncé l'implantation en France de grands centres de recherche en IA, comme le sud-soréen Samsung qui va installer à Paris son troisième centre le plus important au monde. De son côté, le groupe informatique américain IBM a annoncé ce jeudi qu'il allait recruter 400 experts d'intelligence artificielle en France d'ici deux ans.

Document 7

Synthèse du rapport Villani

Posté sur le site de AI for Humanity

Le rapport Villani en 10 pages mars 2018

Partie 1 — Une politique économique articulée autour de la donnée

Les mastodontes actuels de l'intelligence artificielle (États-Unis et Chine) et les pays émergents de la discipline (Israël, Canada et Royaume-Uni notamment) se développent ou se sont développés sur des modèles parfois radicalement différents. Ce n'est pas forcément grâce à un « Google européen » que la France et l'Europe pourront se faire une place sur la scène mondiale de l'IA. Elles doivent pour cela inventer un modèle spécifique.

Un écosystème européen de la donnée

Les données sont généralement le point de départ de toute stratégie en IA, car de leur disponibilité dépendent de nombreux usages et applications. Or les données bénéficient aujourd'hui majoritairement à une poignée de très grands acteurs. Ce n'est qu'au prix d'un plus grand accès et d'une meilleure circulation de ces données, pour en faire bénéficier les pouvoirs publics, mais aussi les acteurs économiques plus petits et la recherche publique, qu'il sera possible de rééquilibrer les rapports de forces.

La puissance publique doit pour cela amorcer de nouveaux modes de production, de collaboration et de gouvernance sur les données, par la constitution de « communs de la donnée »¹. Cela devra passer par une incitation des acteurs économiques au partage et à la mutualisation de leurs données, l'État pouvant ici jouer un rôle de tiers de confiance. Dans certains cas, la puissance publique pourrait imposer l'ouverture

s'agissant de certaines données d'intérêt général. Au niveau européen, plusieurs réformes en cours doivent permettre un meilleur accès et une plus grande circulation des données. La révision prochaine de la directive sur la réutilisation des informations du secteur public doit être l'occasion d'accélérer le mouvement d'ouverture des données publiques et de préciser les modalités d'un accès à des données privées pour des motifs d'intérêt général. La réforme actuelle du droit d'auteur devrait — enfin ! — permettre d'autoriser les pratiques de fouille de texte et de données (text and data mining) dans un objectif de compétitivité de notre recherche publique. Cette politique de la donnée doit s'articuler avec un objectif de souveraineté : la France et l'Europe doivent conserver une position ferme s'agissant du transfert de données hors de l'Union européenne. La stratégie en IA devra en outre capitaliser sur les hauts standards de protection issus de la nouvelle législation européenne sur les données. Le nouveau droit à la portabilité² des individus sur leurs données personnelles pourrait ainsi s'inscrire dans une logique citoyenne, pour permettre à l'État et aux collectivités territoriales de récupérer ces données pour développer des applications en IA à des fins de politique publique.

Renforcer la visibilité de ceux qui font l'IA

La France dispose de tous les atouts pour exister pleinement sur la scène internationale. Néanmoins nos entreprises et nos réseaux académiques souffrent d'un véritable manque de visibilité. C'est à la fois vrai à l'étranger et sur le marché domestique : les grandes entreprises préférant parfois céder aux sirènes des géants mondiaux de la discipline plutôt que de faire confiance à nos pépites nationales, soit parce qu'elles en ignorent l'existence,

¹ Les communs, ou biens communs, désignent une ressource dont l'usage et la gouvernance sont définis par une communauté.

² La capacité pour les utilisateurs de récupérer leurs données, pour leurs propres usages ou pour les transférer vers un autre service.

soit par excès de prudence. Notre mission propose ainsi de fédérer les acteurs français de l'intelligence artificielle autour d'une marque forte, qui pourrait prendre la forme de labels et de prix « d'innovation de terrain » visant à récompenser les solutions d'IA les plus innovantes et à sécuriser de potentiels acheteurs.

Cet effort doit s'accompagner d'une structuration de la demande en IA. Cela pourrait passer par la création d'un guichet unique d'information visant à aider les potentiels acheteurs d'IA à mieux formaliser leurs besoins et à identifier les acteurs permettant d'y répondre.

Un choix clair : mettre l'accent sur quatre secteurs stratégiques

Pour renforcer l'écosystème français et européen de l'intelligence artificielle, il nous faut tirer parti des avantages comparatifs et des niches d'excellence de notre économie. De ce point de vue, notre mission recommande d'éviter les logiques de saupoudrage et de concentrer l'effort sur quatre secteurs prioritaires : santé, environnement, transports-mobilités et défense-sécurité. Tous représentent un défi majeur du point de vue de l'intérêt général, tous requièrent une impulsion importante de l'État et tous sont susceptibles de cristalliser l'intérêt et l'implication continue des acteurs publics et privés.

Pour chacun de ces secteurs, la stratégie industrielle doit permettre de mobiliser et de structurer les écosystèmes autour de grands enjeux et défis sectoriels. Il n'est pas question ici de développer de l'IA pour elle-même, comme une fin en soi, mais justement de canaliser cette énergie pour le développement d'applications, d'usages qui contribuent à améliorer notre performance économique ainsi que le bien commun : détection précoce des

pathologies, médecine des 4^{PS}, disparition des déserts médicaux, mobilité urbaine à zéro émission... Ces enjeux et défis affichés de politique industrielle, propres à chaque secteur, dépassent le sujet de l'IA, mais pourraient contribuer à donner un terrain favorable à son développement.

Le deuxième pilier de cette stratégie consiste à mettre en place des plateformes sectorielles de mutualisation. Celles-ci devront offrir un accès différencié et sécurisé aux acteurs de ces différents écosystèmes (chercheurs, entreprises, puissance publique) à des données pertinentes pour le développement d'IA, à des ressources logicielles ainsi qu'à des infrastructures de calcul d'ampleur significative. Dans un continuum public-privé, ces plateformes devront permettre à ces différents acteurs de développer de nouvelles fonctionnalités adaptées aux spécificités de chaque secteur.

Enfin, il est essentiel de fluidifier les parcours d'innovation en IA avec la mise en place de « bacs à sable d'innovation », qui se déclinent sur trois aspects : un allègement temporaire de certaines contraintes réglementaires pour laisser le champ libre à l'innovation ; un accompagnement des acteurs dans la prise en compte de leurs obligations ; des moyens d'expérimentation en situation réelle.

Transformation de l'État, État exemplaire

L'État doit être un puissant moteur de ces transformations. La puissance publique doit se donner les moyens matériels et humains d'intégrer l'IA à la conduite de ses politiques publiques, à la fois dans une perspective de modernisation et par souci d'exemplarité.

³ Médecine personnalisée, préventive, prédictive, participative.

Le rapport Villani en 10 pages mars 2018

Cette transformation va inévitablement prendre du temps et la maturité des différents ministères et administrations sur l'IA est très inégale. C'est pourquoi il est nécessaire d'installer un coordinateur interministériel dédié à la mise en œuvre de cette stratégie. Celui-ci pourra s'appuyer sur un pôle mutualisé de compétences, constitué d'une trentaine d'agents et chargé de conduire des missions de conseil auprès des administrations.

La commande publique doit être repensée. Évaluée à près de 70 milliards d'euros annuels pour l'État, les établissements publics et les collectivités, elle reste insuffisamment orientée vers l'innovation, pour des raisons à la fois opérationnelles, juridiques et culturelles. Notre mission recommande plusieurs mesures qui visent à mettre l'achat public au service du soutien aux industries européennes et à dynamiser l'achat public innovant.

Partie 2 — Pour une recherche agile et diffusante

La recherche française est au premier plan mondial pour ce qui concerne ses chercheurs en mathématiques et en intelligence artificielle, mais elle a du mal à transformer ses avancées scientifiques en applications industrielles et économiques. Elle pâtit d'une « fuite des cerveaux » vers les géants américains. L'offre de formation se situe par ailleurs bien en deçà des besoins en matière d'IA et science des données.

Fédérer les acteurs de la recherche autour d'Instituts Interdisciplinaires d'Intelligence Artificielle

Il faut renforcer la place mondiale de notre recherche en IA en créant, à l'intérieur d'une sélection d'établissements publics d'enseignement et recherche, des Instituts

Interdisciplinaires d'Intelligence Artificielle (3IA) autonomes et coordonnés, réunissant chercheurs, ingénieurs et étudiants. Ils ont vocation à être répartis géographiquement sur l'ensemble du territoire national, diversifiés thématiquement sur les domaines de l'IA et mettant un accent fort sur l'interdisciplinarité, notamment, mais pas uniquement, vis-à-vis des sciences humaines et sociales.

Il faudra tout d'abord réussir à attirer des scientifiques français et étrangers. Ces instituts devront procurer un environnement de travail suffisamment attractif pour résister à la pression compétitive des géants du numérique, c'est pourquoi ils devront être conçus comme des « zones franches de l'IA » : allègement drastique des formalités administratives du quotidien, compléments de salaire conséquents, aides pour l'amélioration de la qualité de vie. Ces instituts pourront offrir des postes à temps plein et des statuts intermédiaires d'affiliés, pour les chercheurs qui restent dans les établissements fondateurs.

Il faudra également attirer des partenaires privés (grands groupes, PME, startups) porteurs de solutions en IA fondamentalement nouvelles, en leur permettant de former leurs ingénieurs, de recruter des ingénieurs de haut niveau, de réaliser ou consolider des avancées technologiques. Plusieurs modes de participation pourront être définis, selon des contrats-cadres adaptés permettant de démarrer d'éventuelles collaborations spécifiques de manière simple et rapide.

Ces instituts devront effectivement contribuer à l'augmentation substantielle d'offres de formations en IA attractives et diversifiées. La présence d'enseignants de réputation internationale, entourés d'une équipe du même niveau ; la possibilité de côtoyer, via des stages ou des concours d'innovation, des industriels et des acteurs économiques de tout premier plan ; des offres de formations pluridisciplinaires avec

Synthèse

doubles diplômes, et des facilités financières comme des bourses de master et de doctorat devraient permettre d'augmenter significativement le nombre d'étudiants venant se former à l'IA autour de ces instituts.

Il est enfin nécessaire de penser une coordination nationale du réseau des 3IA, sur les plans scientifique et administratif, qui soit efficace et transparente. Au niveau scientifique, cela passe notamment par la coordination des séminaires, le partage des ressources de formation, la coordination des stages et la consolidation de leurs résultats. Au niveau administratif, il s'agit de déterminer les allègements consentis à tous les 3IA, garantissant que chacun d'entre eux puisse en bénéficier, sans pour autant alourdir les procédures et sans empiéter sur l'autonomie de chacun.

Des moyens de calcul pour la recherche

Les instituts de recherche en IA doivent pouvoir disposer d'outils de calcul qui leur permettent de rivaliser avec les moyens quasi-illimités des grands acteurs privés. C'est pourquoi notre mission propose la mise en place d'un supercalculateur conçu spécifiquement pour les applications d'IA, dédié aux chercheurs et à leurs partenaires économiques dans le cadre de projets communs.

Ce supercalculateur, s'il est essentiel, devrait être complété par un forfait d'accès à un cloud privé, développé à un niveau européen et adapté aux spécificités de l'IA (en temps de calcul et en espace de stockage).

Rendre plus attractives les carrières dans la recherche publique

Même s'il est illusoire de penser rivaliser financièrement avec les offres des GAFAM, l'écart est actuellement si important qu'il

tend à décourager les jeunes diplômés, y compris ceux qui sont le plus attachés à la recherche publique et au bien commun. Un doublement des salaires en début de carrière est un point de départ minimal indispensable, sous peine de voir se tarir définitivement le flux de jeunes prêts à s'investir dans l'enseignement supérieur et la recherche académique. Il est par ailleurs nécessaire de renforcer l'attractivité de la France pour les talents expatriés ou étrangers, notamment avec des incitations financières.

Partie 3 — Anticiper les impacts sur le travail, l'emploi et expérimenter

Le monde du travail est à l'aube de grandes transformations et n'y est encore que peu préparé. Les incertitudes sur les conséquences du développement combiné de l'intelligence artificielle, de l'automatisation et de la robotique sont très importantes, notamment concernant les emplois détruits et créés. Néanmoins, il apparaît de plus en plus certain que la majorité des métiers et des organisations vont être très largement transformés. Nous entrons donc dans une période de transition technologique importante : l'histoire nous enseigne que les précédentes transitions ne se sont pas faites sans encombre et que les processus de réajustement politiques ont parfois été violents, souvent au détriment des populations déjà les plus fragiles. Il est donc nécessaire de prendre le problème à bras le corps et d'agir résolument, sans céder à la panique ni au fatalisme.

Cela passe d'abord par une réflexion sur les modes de complémentarité entre l'humain et l'intelligence artificielle : si l'on part du principe que, pour la plupart des métiers, les individus seront amenés à travailler en collaboration avec une machine, il est nécessaire de définir une complémentarité

Le rapport Villani en 10 pages mars 2018

qui ne soit pas aliénante, mais au contraire permette de développer les capacités proprement humaines (créativité, dextérité manuelle, capacité de résolution de problèmes...). Cela pourra s'articuler de plusieurs manières. Tout d'abord au travers d'une transformation du dialogue social afin d'intégrer pleinement les enjeux numériques et développer un indice de bonne complémentarité. De manière plus globale, un chantier législatif relatif aux conditions de travail à l'heure de l'automatisation pourrait être lancé afin de prendre en compte les nouveaux risques. Enfin, une transformation de la formation initiale et continue devra être amorcée afin de favoriser les pédagogies expérimentales, à même de développer les compétences créatives qui deviennent de plus en plus cruciales.

Créer un lab public de la transformation du travail

C'est la première nécessité : s'assurer que la capacité d'anticipation soit pérenne, continue et surtout articulée avec des politiques publiques. La parution des études sur l'avenir du travail occasionne des débats collectifs passionnants, mais souvent sans véritable incidence, tandis que les politiques publiques ne sont modifiées qu'à la marge et peinent à prendre véritablement en compte les résultats de ces exercices prospectifs. Les transformations peuvent être extrêmement rapides et les circuits des politiques publiques sont tout aussi complexes et difficiles à manœuvrer. La formation professionnelle, à elle toute seule par exemple, représente 32 milliards d'euros par an, avec une multitude de canaux de financement et une myriade d'acteurs différents.

Il est donc nécessaire de constituer un espace où les capacités prospectives, de prévisions macro-économiques et d'analyse des mutations des usages puissent être mises en lien avec des capacités d'expérimentation concrètes et articulées

avec des actions à destination de certaines catégories de travailleurs. Une structure pérenne pourrait donc être installée, qui aurait un rôle de « tête chercheuse » à l'intérieur des politiques publiques de l'emploi et de la formation professionnelle. Elle aura un double rôle : anticiper et expérimenter.

La démarche expérimentale pourra servir à amorcer des logiques différentes de celles qui sont actuellement en vigueur dans la formation professionnelle. Les dispositifs actuels sont largement « à la main » des salariés, dans une logique de responsabilisation individuelle. Au vu du caractère potentiellement très rapide, voire exponentiel de ces transformations, il semble difficile, pour les dispositifs généraux existants, de répondre à l'ensemble des situations et de permettre à la fois la prise en compte des besoins de l'ensemble de la population et la nécessité d'agir de manière ciblée et urgente. De plus, face à la transformation de leur emploi, les individus ne sont pas égaux dans la capacité de s'adapter et de construire des parcours professionnels.

À cet égard, des expérimentations pourraient être menées afin de construire des dispositifs qui ciblent certaines populations d'individus, dont les emplois sont considérés comme étant le plus à risque d'automatisation et pour lesquelles il sera complexe d'amorcer seules leur transition professionnelle. Il s'agit donc de rompre, en partie, avec la seule logique de responsabilisation de l'individu concernant sa propre transition professionnelle.

Expérimenter de nouveaux modes de financement de la formation professionnelle pour tenir compte des transferts de valeur

Le financement de la formation professionnelle est fondé sur la masse salariale. Or, le développement de l'IA

renforce la mutation des chaînes de valeur et entraîne une décomposition entre les acteurs qui financent la formation professionnelle et ceux qui captent la valeur ajoutée. Ainsi des acteurs ayant une très faible masse salariale peuvent être à l'origine d'une grande partie de la valeur ajoutée d'une chaîne de valeur globale qu'ils contribuent à très largement modifier par exemple en développant un logiciel pour les voitures autonomes. Pour autant, à l'heure actuelle, ils ne participent pas au financement de la transition professionnelle des individus employés par d'autres acteurs de la chaîne de valeur.

Il est donc proposé d'instaurer un dialogue social autour du partage de la valeur ajoutée au niveau de la chaîne de valeur entière. Ce type de négociation ne correspond pas aux structurations habituelles du dialogue social qui fonctionne très largement à un niveau national et surtout suivant une structuration verticale, par branche. Des expérimentations pourraient être organisées par l'Organisation internationale du travail, ou encore les comités de dialogue social sectoriel, autour de produits et de chaînes de valeur particulièrement symptomatiques des phénomènes de captation de valeur.

Former des talents en IA, à tous niveaux

Un objectif clair doit être fixé : à horizon trois ans, multiplier par trois le nombre de personnes formées en intelligence artificielle en France, à la fois en faisant en sorte que l'offre de formation existante s'oriente vers l'IA, mais également en créant de nouveaux cursus et de nouvelles formations à l'IA (doubles cursus droit-IA par exemple, modules généraux...). L'ensemble des niveaux (bac +2, bac +3, master, doctorat) doit faire l'objet d'attention.

Partie 4 — L'intelligence artificielle au service d'une économie plus écologique

Donner un sens à l'intelligence artificielle, c'est également penser sa soutenabilité, notamment écologique. Cela ne se résume pas à lister les usages de l'IA qui pourront aider à la transition écologique. Il s'agit de penser une IA nativement écologique et de l'utiliser pour mieux penser l'impact de l'humain sur son environnement. Il y a urgence : d'ici 2040, les besoins en espace de stockage au niveau mondial, fondamentalement corrélés au développement du numérique et de l'IA, risquent d'excéder la production disponible globale de silicium.

La France et l'Europe peuvent devenir le fer-de-lance de cette transition écologique intelligente, d'abord en inscrivant le sujet à l'agenda international. Premier chantier : penser les impacts de l'IA sur la réalisation des objectifs de l'ONU sur le développement durable (ODD) — en quoi celle-ci en met certains sous contrainte, comment elle peut à l'inverse permettre d'en accélérer d'autres. L'IA doit s'intégrer aux initiatives émergentes dans le cadre de l'Accord climat et du Pacte mondial pour l'environnement.

Les acteurs des transitions numérique et écologique doivent se fédérer. Pour cela, il est nécessaire de créer un lieu dédié à cette rencontre entre la recherche en IA et la recherche portant sur l'optimisation des ressources énergétiques. Il s'agira de porter des projets à la croisée des sciences du vivant et de l'écologie, la recherche sur le climat et la météo.

Le consommateur doit être acteur dans le verdissement de ces technologies. Notre mission propose ainsi la mise en place d'une plateforme dédiée à la mesure de l'impact environnemental des solutions numériques

Le rapport Villani en 10 pages mars 2018

intelligentes. Cette plateforme devra s'accompagner d'un outil simple permettant à tout citoyen de prendre conscience de ces enjeux et de comparer l'impact environnemental de ces différents produits et services, logiciels et matériels.

Penser une IA plus verte

Nous devons penser l'innovation de rupture dans le domaine du semi-conducteur, l'une des briques matérielles de l'IA. À ce titre les technologies neuromorphiques⁴ peuvent permettre des économies d'énergie considérables — et la France est déjà très avancée dans le domaine.

Par ailleurs, les pouvoirs publics doivent agir pour le verdissement de la chaîne de valeur et accompagner l'industrie du cloud européen dans le sens de sa transition écologique. Certains acteurs sont déjà exemplaires en matière d'optimisation de l'utilisation de l'énergie. Il est important de diffuser ces bonnes pratiques à l'ensemble du secteur. Un label pourrait être mis en place afin de valoriser les solutions les plus exemplaires.

Enfin, le verdissement de la chaîne de valeur de l'IA passera nécessairement par des architectures matérielles et logicielles ouvertes (open hardware et open software) qui, en plus d'être un facteur de confiance, peuvent permettre des économies d'énergie significatives et qui peuvent inspirer les initiatives en cours au niveau européen.

Libérer la donnée écologique

Le développement d'une IA verte n'est possible qu'à condition de libérer la donnée écologique. Pour développer des solutions d'IA au service de la transition écologique, il est ainsi primordial de mettre à la

disposition de tous, chercheurs et entreprises européennes, et rapidement, à horizon 2019, les données publiques disponibles : météorologiques, agricoles, de transports, d'énergie, de biodiversité, de climat, de déchets, cadastrales, de diagnostic de performance énergétique... Pour les jeux de données les plus sensibles, l'ouverture pourrait se faire dans un périmètre précis, par exemple dans le cadre de défis sectoriels. Il est également essentiel de libérer la donnée privée, lorsque cela est nécessaire.

Partie 5 — Quelle éthique de l'IA ?

Les progrès récents de l'IA dans de nombreux domaines (voitures autonomes, reconnaissance d'images, assistants virtuels) et son influence croissante sur nos vies renforcent sa place dans le débat public. Ce débat a notamment pris la forme d'une large réflexion sur les enjeux éthiques liés au développement des technologies d'intelligence artificielle et plus largement des algorithmes. Loin des considérations spéculatives sur les menaces existentielles de l'IA pour l'humanité, les réflexions tendent à se cristalliser autour des algorithmes du « quotidien », qui peuvent d'ores et déjà avoir des conséquences importantes sur nos vies.

Si nous souhaitons faire émerger des technologies d'IA conformes à nos valeurs et normes sociales, il faut agir dès à présent en mobilisant la communauté scientifique, les pouvoirs publics, les industriels, les entrepreneurs et les organisations de la société civile. Notre mission a cherché, humblement, à proposer quelques pistes permettant de poser les bases d'un cadre éthique pour le développement de l'IA et à faire vivre ce débat dans la société.

⁴ On appelle puces neuromorphiques les puces dont le fonctionnement s'inspire du cerveau humain.

Ouvrir les boîtes noires

Une grande partie des considérations éthiques soulevées tiennent à l'opacité de ces technologies : l'IA donne aujourd'hui des résultats spectaculaires, pour des raisons que les chercheurs ont parfois du mal à expliquer. C'est le fameux problème de la boîte noire : des systèmes algorithmiques dont il est possible d'observer les données d'entrée (input), les données de sortie (output), mais dont on comprend mal le fonctionnement interne. Dans un contexte où l'IA est susceptible de reproduire des biais et des discriminations, et à mesure de son intrusion dans nos vies sociales et économiques, être en mesure « d'ouvrir les boîtes noires » tient de l'enjeu démocratique.

L'explicabilité des algorithmes d'apprentissage automatique est un sujet si pressant qu'il constitue aujourd'hui un champ de recherche spécifique, qui doit être soutenu par la puissance publique. Trois axes en particulier semblent mériter une attention particulière : la production de modèles plus explicables bien sûr, mais aussi la production d'interfaces utilisateurs plus intelligibles et la compréhension des mécanismes cognitifs à l'œuvre pour produire une explication satisfaisante.

Au-delà de la transparence, il est nécessaire d'accroître l'auditabilité des systèmes d'IA. Cela pourrait passer par la constitution d'un corps d'experts publics assermentés, en mesure de procéder à des audits d'algorithmes, des bases de données et de procéder à des tests par tout moyen requis. Ces experts pourraient être saisis à l'occasion d'un contentieux judiciaire, dans le cadre d'une enquête diligentée par une autorité administrative indépendante ou suite à une demande du Défenseur des droits.

Penser l'éthique dès la conception

Les chercheurs, ingénieurs et entrepreneurs qui contribuent à la conception, au développement et à la commercialisation de systèmes d'IA sont amenés à jouer un rôle décisif dans la société numérique de demain. Il est essentiel qu'ils agissent de manière responsable, en prenant en considération les impacts socio-économiques de leurs activités. Pour s'en assurer, il est nécessaire de les sensibiliser, dès le début de leur formation, aux enjeux éthiques liés au développement des technologies numériques. Aujourd'hui cet enseignement est quasiment absent des cursus des écoles d'ingénieurs ou des parcours informatiques des universités, alors même que le volume et la complexité des problématiques éthiques auxquels des futurs diplômés seront confrontés ne cessent de croître.

Au-delà de la formation des ingénieurs, les considérations éthiques doivent irriguer le développement même des algorithmes d'intelligence artificielle. Sur le modèle de l'étude d'impact sur les risques en matière de vie privée (*privacy impact assessment*), rendu obligatoire pour certains traitements de données par le règlement général sur la protection des données (RGPD), il pourrait être institué une étude d'impact sur les risques de discrimination (*discrimination impact assessment*). L'objectif est simple : obliger les développeurs d'IA à se poser les bonnes questions, au bon moment.

Plus généralement, l'utilisation croissante de l'IA dans certains domaines sensibles comme la police, la banque, l'assurance, la justice ou l'armée (avec la question des armes autonomes) appelle un véritable débat de société et une réflexion sur la question de la responsabilité humaine. Nous devons également nous interroger sur la place de l'automatisation dans les décisions humaines : existent-ils des domaines où le jugement humain, aussi

Le rapport Villani en 10 pages mars 2018

faillible soit-il, ne devrait pas à être remplacé par une machine ?

Créer un comité d'éthique de l'IA

Notre mission recommande la création d'un comité d'éthique des technologies numériques et de l'IA ouvert sur la société. Cet organe serait chargé d'organiser le débat public, de façon lisible, construite et encadrée par la loi. Il devra parvenir à articuler des logiques de temps court, celui des enjeux économiques et industriels, en bonne interaction avec les comités sectoriels, tout en parvenant à s'en extraire pour penser le temps long.

Les avis de ce comité, élaborés en toute indépendance, pourraient éclairer les choix technologiques des chercheurs, des acteurs économiques, industriels et de l'État. Ses recommandations pourront servir de référence pour la résolution de dilemmes éthiques (par exemple sur le véhicule autonome) et donc servir de standard pour les développements en IA.

Partie 6 — Pour une IA inclusive et diverse

L'intelligence artificielle ne peut pas être une nouvelle machine à exclusion. C'est une exigence démocratique dans un contexte où ces technologies sont en passe de devenir une des clés du monde à venir. Elle ouvre de formidables opportunités de création de valeur et de développement de nos sociétés et des individus. Ces opportunités doivent bénéficier à tous.

Mixité et diversité : agir pour l'égalité

En dépit d'une féminisation lente, mais progressive des filières scientifiques et techniques, le numérique fait figure d'exception : la parité entre les hommes et

les femmes est loin d'y être acquise. À mesure que le numérique et, demain, l'intelligence artificielle deviennent omniprésents dans nos vies, ce manque de diversité peut conduire les algorithmes à reproduire des biais cognitifs — souvent inconscients — dans la conception des programmes, l'analyse des données et l'interprétation des résultats. L'un des grands défis de l'IA consiste donc à parvenir à une meilleure représentativité de nos sociétés.

Si l'éducation à l'égalité et au numérique est une condition préalable et essentielle, la mixité pourrait être atteinte avec une politique incitative visant à atteindre un seuil de 40% d'étudiantes dans les filières du numérique (classes préparatoires et filières des grandes écoles et des universités) d'ici 2020.

L'ensemble des actions en faveur de la diversité dans les entreprises du numérique pourraient par ailleurs être portées par une action nationale en faveur de la mixité et de la diversité dans la technologie avec l'alimentation, notamment, d'une base de données nationale permettant d'objectiver les inégalités entre les femmes et les hommes au travail et de fonds dédiés à soutenir la diversité dans l'IA.

Développer la médiation numérique et l'innovation sociale pour que l'IA bénéficie à tous

Face à l'ampleur des transformations à venir par l'IA, il est de notre responsabilité collective de s'assurer que personne ne soit mis de côté. Pour que chacun puisse véritablement bénéficier des avancées de l'IA, nos procédures d'accès aux droits doivent évoluer, et nos capacités de médiation considérablement se renforcer. Notre mission propose donc de mettre en place un système automatisé d'aide à la gestion des démarches administratives qui vise à améliorer la connaissance pour le grand public des règles administratives et

Synthèse

de leur application à une situation personnelle. En complément, de nouvelles capacités de médiation doivent être déployées pour accompagner les personnes qui en ont besoin, en lien avec les réseaux de prise en charge présents sur le territoire.

Enfin, il est important que la puissance publique soutienne le développement d'initiatives basées sur l'IA dans les champs sociaux. À l'heure actuelle, les capacités d'innovation par l'IA restent très concentrées au sein d'un petit nombre d'entreprises. À l'exception de la santé, les champs sociaux reçoivent une part minoritaire des investissements privés. Cette structuration de l'écosystème d'innovation en IA a des conséquences sur la vitesse des progrès réalisés dans les champs sociaux. Afin de redistribuer ces capacités d'innovation, la puissance publique pourrait lancer des programmes spécifiques pour accompagner l'innovation d'IA en matière sociale, et outiller les acteurs sociaux afin qu'ils puissent bénéficier des avancées liées à l'IA.

Document 8

CNIL Extrait du rapport « Comment permettre à l'homme de garder la main - Les enjeux éthiques des algorithmes et de l'intelligence artificielle »

15 décembre 2017

La CNIL publie le rapport de synthèse du débat public qu'elle a animé sur les enjeux éthiques des algorithmes et de l'intelligence artificielle. Isabelle FALQUE-PIERROTIN, présidente de la CNIL, le présente aujourd'hui en présence de Mounir MAHJoubi, Secrétaire d'Etat chargé du Numérique et de Cédric VILLANI, Député, chargé par le gouvernement d'une mission sur l'intelligence artificielle.



L'irruption dans nos vies quotidiennes des algorithmes et de l'intelligence artificielle fait l'objet d'une attention publique soutenue, depuis que quelques cas emblématiques (APB, rôle des réseaux sociaux dans les présidentielles américaines etc.) ont marqué les citoyens, ici en France ou à l'étranger. A l'heure où se construisent les positions française et européenne sur l'intelligence artificielle, ce rapport a pour but de construire un panorama des enjeux éthiques qu'ils soulèvent.

Un débat public innovant sur les algorithmes et l'intelligence artificielle

Dans notre monde de plus en plus numérique, les algorithmes et l'intelligence artificielle sont partout : pour simuler l'évolution de la propagation de la grippe en hiver, pour recommander des livres à des clients, pour suggérer aux forces de police des zones où patrouiller en priorité, pour piloter de façon autonome des automobiles, pour élaborer automatiquement un diagnostic médical personnalisé, pour personnaliser un fil d'activité sur les réseaux sociaux etc. Des tâches complexes, parfois critiques, sont ainsi déléguées à des systèmes de plus en plus autonomes à mesure que les techniques d'apprentissage propres à l'intelligence artificielle se développent. Ces usages multiples et croissants, indissociables des masses de données inédites à disposition dans tous les secteurs, soulèvent de nombreuses questions.

Chargée par loi pour une République numérique de mener une réflexion sur les questions éthiques et de société posées par les nouvelles technologies, la CNIL s'est naturellement tournée vers le thème des **algorithmes à l'heure de l'intelligence artificielle**.

Le rapport s'appuie sur les résultats d'un débat public ouvert et décentralisé que la CNIL a animé de janvier à octobre 2017 : 3 000 personnes ont participé à 45 manifestations qui se sont tenues à l'initiative de 60 partenaires en France (Ax les Termes, Bordeaux, Caen, Lille, Lyon, Marseille, Paris, Toulouse) et à l'étranger (Etats-Unis). Une concertation citoyenne a également été organisée à Montpellier le 14 octobre pour compléter la démarche.

Médias, vie publique et politique, éducation, culture, santé, justice, sécurité et défense, banque et finance, emploi et recrutement ont été abordés comme champs d'exploration des algorithmes et de l'intelligence artificielle par les partenaires : des instituts de recherches (CREOGN, CNAM, ENSC, INSA), des institutions publiques (ministères, universités et grandes écoles), des fédérations professionnelles, des syndicats, des ordres d'avocat, des hôpitaux, des think tanks, des entreprises, etc.

Pour Isabelle Falque-Pierrotin, Présidente de la CNIL :

L'objectif de ce débat est de garantir que l'intelligence artificielle augmente l'homme plutôt qu'elle ne le supprime et participe à l'élaboration d'un modèle français de gouvernance éthique de l'intelligence artificielle. Nous devons collectivement faire en sorte que ces nouveaux outils soient à la main humaine, à son service, dans un rapport de transparence et de responsabilité

Quelles réponses éthiques au développement des algorithmes et de l'intelligence artificielle ?

Vers l'affirmation de deux principes fondateurs : loyauté et vigilance

Les débats ont permis de dégager deux principes fondateurs pour une intelligence artificielle au service de l'homme. Ces principes pourraient s'inscrire dans **une nouvelle génération de garanties et de droits fondamentaux à l'ère numérique**, des « droits-système » organisant la **gouvernance mondiale de notre univers numérique** :

- Un principe de loyauté appliqué à tous les algorithmes et intégrant les impacts collectifs, et pas seulement personnels, de ces derniers. Tout algorithme, qu'il traite ou non des données personnelles, doit être loyal envers ses utilisateurs, non pas seulement en tant que consommateurs, mais également en tant que citoyens, voire envers des communautés ou de grands intérêts collectifs dont l'existence pourrait être directement affectée. **L'intérêt des utilisateurs doit primer**. Par exemple, un tel principe pourrait avoir vocation à s'appliquer à l'impact potentiel des réseaux sociaux sur la structure du débat public dans nos démocraties (segmentation du corps politique par le ciblage de l'information) ou à celui d'algorithmes de police prédictive sur des communautés ou quartiers entiers.
- Un principe de vigilance/réflexivité : il s'agit d'organiser une forme de questionnement régulier, méthodique et délibératif à l'égard de ces objets mouvants. Ce principe constitue une réponse directe aux exigences qu'imposent ces objets technologiques du fait de leur nature imprévisible (inhérente au *machine learning*), du caractère très compartimenté des chaînes

algorithmiques au sein desquels ils s'insèrent et, enfin, de la confiance excessive à laquelle ils donnent souvent lieu. C'est l'ensemble des maillons de la chaîne algorithmique (concepteurs, entreprises, citoyens) qui doivent être mobilisés pour donner corps à ce principe, au moyen de procédures concrètes (par exemple, des comités d'éthique assurant un dialogue systématique et continu entre les différentes parties-prenantes).

Ces principes fondateurs sont complétés par des principes organisationnels ayant trait à l'intelligibilité et à la responsabilité des systèmes algorithmiques ainsi qu'à la nature de l'intervention humaine dans la prise de décision algorithmique.

Les 6 recommandations opérationnelles

Ces principes font l'objet d'une déclinaison sous la forme de **6 recommandations opérationnelles** à destination tant des pouvoirs publics que des diverses composantes de la société civile (entreprises, grand public, etc.) :

1. Former à l'éthique tous les acteurs-maillons de la « chaîne algorithmique » (concepteurs, professionnels, citoyens) : l'alphabétisation au numérique doit permettre à chaque humain de comprendre les ressorts de la machine ;
2. Rendre les systèmes algorithmiques compréhensibles en renforçant les droits existants et en organisant la médiation avec les utilisateurs ;
3. Travailler le *design* des systèmes algorithmiques au service de la liberté humaine, pour contrer l'effet « boîtes noires » ;
4. Constituer une plateforme nationale d'audit des algorithmes ;
5. Encourager la recherche sur l'IA éthique et lancer une grande cause nationale participative autour d'un projet de recherche d'intérêt général ;
6. Renforcer la fonction éthique au sein des entreprises (par exemple, l'élaboration de comités d'éthique, la diffusion de bonnes pratiques sectorielles ou la révision de chartes de déontologie peuvent être envisagées).

Les problématiques soulevées par les algorithmes et l'IA

Pour formuler ces principes fondateurs et les recommandations, le rapport s'est basé sur 6 problématiques essentielles remontées lors des débats afin de relever le défi éthique posé par ces technologies :

1. L'autonomie humaine au défi de l'autonomie des machines :

Algorithmes et intelligence artificielle permettent la délégation croissante de tâches, de raisonnements ou de décisions de plus en plus critiques à des machines. Souvent jugées infaillibles et « neutres », celles-ci n'ouvrent-elles pas la voie à une confiance excessive et à la tentation pour chacun de ne pas exercer pleinement ses responsabilités ? Comment faire face aux formes nouvelles de dilution de la responsabilité qu'impliquent des systèmes algorithmiques complexes et très segmentés ?

2. Biais, discrimination et exclusion :

Biais, discrimination et exclusion constituent des effets clairement identifiés des algorithmes et de l'intelligence artificielle. Parfois volontaires, ces phénomènes sont, à l'heure du déploiement des algorithmes de *machine learning*, le plus souvent inconscients et difficilement repérables. Comment appréhender cette nouvelle problématique ?

3. Fragmentation algorithmique : la personnalisation contre les logiques collectives :

La personnalisation constitue l'une des grandes promesses de l'algorithme tel qu'il s'est construit avec le Web. Bien que des bénéfices certains pour l'individu émanent de la segmentation et du profilage, comment peuvent-ils tout autant affecter vigoureusement des logiques collectives essentielles à la vie de nos sociétés (pluralisme démocratique et culturel, mutualisation du risque) ?

4. Entre limitation des mégafichiers et développement de l'intelligence artificielle : un équilibre à réinventer :

L'intelligence artificielle nécessite d'importantes quantités de données à des fins d'apprentissage. Elle suscite également de nombreuses promesses. Comment exploiter celles-ci tout en maintenant l'impératif de protéger les libertés individuelles de chacun, intrinsèque à la législation européenne de protection des données personnelles ?

5. Qualité, quantité, pertinence : l'enjeu des données fournies à l'IA :

L'entraînement d'un algorithme d'apprentissage implique de l'alimenter avec des données sélectionnées avec soin, pertinentes au regard de l'objectif poursuivi et en quantité suffisante. Comment parvenir à une telle attitude critique en dépit de la tendance à une confiance excessive dans la machine ?

6. L'identité humaine au défi de l'intelligence artificielle :

L'idée d'une spécificité humaine irréductible se voit questionnée par l'autonomie grandissante des machines ainsi que par l'apparition de formes d'hybridation entre humains et machines. Faut-il et est-il possible de parler au sens propre d'« éthique des algorithmes » ? Comment appréhender cette nouvelle classe d'objets que sont les robots humanoïdes, susceptibles d'engendrer chez l'homme des formes d'affect ?

Document n°9

Le Monde Une intelligence artificielle peut-elle devenir présidente des Etats-Unis ?

Publié le 17 février 2016

Il ressemble à tous les autres sites des candidats à la présidentielle américaine. Un drapeau étoilé flotte au vent, le design est soigné, agrémenté de vidéos, le propos est clair, argumenté, direct. A la différence près que ce candidat-là n'est pas humain : il s'agit de Watson, le programme d'intelligence artificielle phare d'IBM, l'un des plus avancés au monde.

« *Nous pensons que les capacités uniques de Watson pour analyser l'information, et prendre des décisions éclairées et transparentes, en font un candidat idéal pour le poste à responsabilités que représente celui de président* », peut-on lire sur le site de la campagne Watson 2016. Sur un ton on ne peut plus sérieux, le site déroule son argumentaire :

« Plus Watson intègre d'informations, plus ses capacités de prise de décision sont efficaces. Il est capable d'analyser des informations venant de n'importe quelle source, il peut donc prendre en compte différentes perspectives et opinions sur tous les sujets. [...] C'est une tâche que doivent effectuer quotidiennement les politiques, y compris le président, et qui pourrait être effectuée de façon plus appropriée et efficace par une intelligence artificielle. »

Selon ce site, Watson pourrait analyser, en prenant en compte de nombreux paramètres, les qualités et défauts de chaque décision, en évaluant « *son impact sur l'économie, l'environnement, l'éducation, la santé, la diplomatie et les libertés publiques* ».

« Frustration et désillusion »

D'où vient ce site ? S'agit-il d'une campagne de communication d'IBM ? L'entreprise a répondu au *Monde* qu'elle n'avait rien à voir avec ce projet, et refuse de le commenter. Peut-être vient-il alors de l'Electronic Frontier Foundation, une importante association de défense des libertés numériques, à laquelle le site propose de faire des dons ? Non plus.

Cette initiative vient en réalité de l'artiste et designer Aaron Siegel, professeur de design à l'université de Californie du Sud, qui se présente comme le « *directeur de campagne* » de Watson 2016. Comme il l'explique au *Monde*, ce projet est issu « *de la frustration et de la désillusion vis-à-vis du processus de l'élection présidentielle américaine* ». Il fustige la façon dont les candidats s'en tiennent à la vision de leur parti « *au lieu d'aborder les problèmes de façon objective* », le coût démesuré des campagnes qui, selon lui, oblige les candidats à se soumettre aux puissances de l'argent, et évoque un « *besoin de transparence en politique* ».

« *Je me suis demandé quelle personne pouvait être le politicien le plus objectif, efficace et non partisan, et je me suis rendu compte que cette personne était un ordinateur.* » Dans son rôle de directeur de campagne, il assure que Watson représente une solution viable :

« Nous pensons qu'une intelligence artificielle telle que Watson peut apporter les capacités de prise de décision objectives dont nous avons besoin chez un dirigeant, ainsi que la transparence nécessaire pour analyser comment ces décisions ont été prises et pourquoi. Le système n'est lié à aucun parti, ses décisions sont donc fondées uniquement sur l'information dont il dispose, et non sur des idéologies. »

« Réponses immédiates à une large gamme de sujets »

Est-il vraiment sérieux ? « *Au vu des problèmes listés sur le site de la campagne, je suis assez sérieux* », assure-t-il, tout en précisant que « *le but de cette campagne est de présenter une alternative à la façon dont le gouvernement fonctionne, pour faire en sorte que les gens réagissent à cette idée.* »

Voilà donc l'objectif réel de Watson 2016 : interroger le système politique actuel, mais aussi la place de l'intelligence artificielle dans notre monde. « *J'espère que cela poussera les gens à discuter du potentiel de l'intelligence artificielle dans la politique.* »

Aujourd'hui, le programme d'IBM est déjà utilisé à des fins médicales, puisque, en analysant les données d'un patient, elle est capable d'aider les médecins à établir un diagnostic. Sur le site d'IBM, l'entreprise vante aussi l'utilité de Watson dans le secteur public :

« Il est parfois difficile pour les organismes du secteur public de répondre aux questions qui leur sont posées, et ce, du fait de la difficulté à s'y retrouver dans l'immense masse de données qu'ils possèdent. Les capacités analytiques de Watson peuvent permettre d'apporter une réponse immédiate à des questions touchant une large gamme de sujets : *“Quelles sont les règles de plan d'occupation des sols pour construire un porche ?” ; “Cette taxe s'applique-t-elle à moi ?” ; “Quelle est la meilleure façon d'obtenir un visa ?”* »

« Rendre Watson open source »

On est encore loin de la capacité de prise de décision politique fantasmée par Aaron Siegel, malgré les prouesses dont est déjà capable son candidat. Watson avait notamment occupé le devant de la scène en 2011 en battant ses adversaires humains au jeu télévisé américain « Jeopardy! ». Dans ce classique de la télévision américaine, le présentateur lit une réponse et les participants doivent deviner la question qui s'y rapporte. Après deux jours de compétition, Watson avait écrasé ses concurrents — une étape symbolique dans l'histoire de l'intelligence artificielle. « *Si vous voulez devenir président, il vaut mieux être déjà un peu connu* », plaisante Aaron Siegel quand on lui demande pourquoi il a choisi ce programme plutôt qu'un autre.

Quand bien même Watson serait un jour capable de prendre des décisions politiques, s'agirait-il réellement de décisions objectives ? La façon dont est codé un programme serait-elle vraiment exempte de tout biais politique ?

« *C'est exactement le type de questions que j'espérais que cela soulève*, répond Aaron Siegel. *Un des principaux problèmes que les gens ont soulevés concernant Watson est qu'il est développé par IBM. Ne pas savoir exactement comment il fonctionne les met mal à l'aise. Cela pourrait être réglé en rendant Watson open source* », c'est-à-dire en rendant son code source accessible à tous, et donc transparent.

Document n°10

France 24 « L'intelligence artificielle inquiète les grands cerveaux contemporains »

Publié le : 02/02/2015

Plusieurs personnalités de renom, comme le multimilliardaire Bill Gates et le très médiatique physicien Stephen Hawking, mettent en garde contre les dangers de l'intelligence artificielle, une avancée qui menacerait l'humanité.

Qu'ont en commun Bill Gates, l'homme le plus riche au monde, Stephen Hawking, le plus célèbre des physiciens et Elon Musk, le roi des voitures électriques et pionnier du tourisme spatial ? Ils craignent tous que l'intelligence artificielle (IA) se transforme en menace pour l'homme.

"Je suis de ceux qui redoutent la montée en puissance des superintelligences et je ne comprends pas qu'on puisse ne pas être inquiet", a ainsi affirmé Bill Gates lors d'un échange avec des internautes sur le site Reddit, jeudi 29 janvier. Le fondateur de Microsoft et inventeur de Windows rejoint ainsi un cortège de plus en plus peuplé de personnalités qui mettent en garde contre la montée en puissance des IA.

Fin de l'humanité ?

Stephen Hawking a ouvert le bal des hostilités en novembre 2014. Sur un ton quasi-apocalyptique, il avait assuré à la chaîne britannique BBC que "le développement d'une intelligence artificielle pourrait mettre fin à l'humanité". Clive Sinclair, inventeur britannique et pionnier de l'informatique, voit également l'avenir en noir. "À partir du moment où vous développez des machines qui rivalisent avec l'homme en termes d'intelligence, nous allons avoir du mal à survivre", a-t-il affirmé à la BBC.

Quant à Elon Musk, pour ne pas être en reste dans la guerre des prédictions anxiogènes, il a assuré sur Twitter que l'IA était aussi dangereuse, sinon plus, que "les armes atomiques". Même l'institut américain Future of Life, pourtant ouvert aux innovations technologiques et scientifiques, a publié, mi-janvier, une lettre ouverte cosignée par plusieurs centaines de scientifiques qui appellent à une plus grande vigilance dans la recherche sur l'intelligence artificielle.

Les Terminator - ces robots qui ont pris le pouvoir dans le film éponyme de James Cameron - ne sont pourtant pas aux portes des villes. La montée au créneau des scientifiques et autres pontes des nouvelles technologies vient de ce qu'ils sont les plus exposés "à la prédominance du web dans la vie quotidienne et à un monde qui devient de plus en plus connecté donc vulnérable", analyse Peter Ford Dominey, chercheur au CNRS spécialiste des robots et de l'étude du cerveau.

Quand les robots traders dérapent

"Ce qui a, aussi, mis la puce à l'oreille de beaucoup, c'est le travail des robots traders",

ajoute Jean-Michel Besnier, philosophe et chercheur au CNRS spécialiste des problématiques liées à l'intelligence artificielle. Les transactions financières par algorithme assisté - le fondement du trading à haute fréquence - ont déjà causé en mai 2010 un crash à la bourse de New York. Pour Peter Ford Dominey, "ces formes assez simples d'intelligence artificielle peuvent entraîner des problèmes sérieux" capables d'affecter l'économie réelle.

Qu'advient-il lorsque l'intelligence artificielle sera en mesure de prendre des initiatives plus complexes que de décider d'acheter ou vendre des actions ? L'idée d'un système entièrement automatisé qui gèrerait les infrastructures critiques d'un pays a, pour ces chercheurs, de quoi faire froid dans le dos. "Je souscris pleinement à l'appel à la vigilance des scientifiques et autres personnalités s'il s'agit de mettre en garde contre une dominance de l'algorithme [dans le processus de prise de décision, NDLR]", précise Jean-Michel Besnier.

Le spectre d'une "intelligence non-biologique" - comme l'ordinateur de bord du vaisseau dans "2001 : l'Odyssée de l'espace" - qui reléguerait l'homme au second plan relève davantage de la science-fiction pour ces scientifiques.

Les vrais Terminator seraient plutôt des petits bouts de code qui, sous prétexte de simplifier la vie de tout un chacun, priverait l'homme de prises d'initiative. En ce sens, comme le souligne le quotidien "Les Echos", les sorties de Bill Gates ou d'Elon Musk pourraient être autant de critiques dissimulées de Google. Entre son initiative en faveur des voitures sans conducteur et sa frénésie de rachat des sociétés spécialisées dans l'IA (Dark Blue Labs et Vision Factory), le géant de l'Internet paverait l'enfer de bonnes intentions numériques.

Remède à la pauvreté

Mais ce n'est pas, pour autant, une raison pour mettre le holà à toute recherche dans ce domaine, jugent ces spécialistes. "Je ne suis pas de ceux qui croient que la technologie a tendance à s'auto-améliorer toute seule et à croître sans contrôle possible", assure Jean-Michel Besnier. Il suffit de mettre en place les garde-fous à temps.

"Toute technologie à risque, comme le nucléaire ou les OGM, sont soumises à une réflexion sérieuse menée par les politiques et la société civile", rappelle le chercheur Peter Ford Dominey. Il doit en être de même, d'après lui, pour l'intelligence artificielle.

Reste que l'IA, contrairement à l'arme nucléaire, n'est pas intrinsèquement dangereuse. "Le risque provient des responsabilités que nous donnons aux machines", précise Peter Ford Dominey. Dans sa lettre ouverte, l'institut Future for Life souligne d'ailleurs que le risque peut être une chance si la recherche est bien orientée. "Les bénéfices sont énormes puisque l'intelligence artificielle pourrait permettre de trouver des solutions à la pauvreté et mettre au point des remèdes pour la plupart des maladies", assurent les scientifiques qui ont rédigé ce texte. Après tout, si Stephen Hawking a pu lancer son appel à la vigilance, c'est en partie grâce à un ordinateur très avancé qui permet à ce physicien atteint de la maladie de Charcot de s'exprimer.

Document 11

Encyclopédie Larousse « Définition de l'intelligence artificielle »

Intelligence artificielle

Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine.

Avec l'intelligence artificielle, l'homme côtoie un de ses rêves prométhéens les plus ambitieux : fabriquer des machines dotées d'un « esprit » semblable au sien. Pour John MacCarthy, l'un des créateurs de ce concept, « toute activité intellectuelle peut être décrite avec suffisamment de précision pour être simulée par une machine ». Tel est le pari – au demeurant très controversé au sein même de la discipline – de ces chercheurs à la croisée de l'informatique, de l'électronique et des sciences cognitives.

Malgré les débats fondamentaux qu'elle suscite, l'intelligence artificielle a produit nombre de réalisations spectaculaires, par exemple dans les domaines de la reconnaissance des formes ou de la voix, de l'aide à la décision ou de la robotique.