



Opinions Libres

le blog d'Olivier Ezratty

Chronique d'une semaine avec et sans IA

On voit toujours midi à sa porte. Je vais m'adonner ici à l'exercice et parcourir la semaine qui se termine au travers du prisme de l'IA qui est l'un des principaux centres de mon activité du moment. Les outils à base d'IA se sont insérés dans la vie de tous les jours, ne serait-ce que dans nos smartphones pour leur fonction de mise au point photo ou pour la commande vocale. Ils se cachent aussi de manière plus ou moins visible dans l'actualité quotidienne.

L'IA n'est évidemment pas une solution miracle à tous les problèmes humains. C'est toutefois une riche boîte à outils exploitable dans de nombreux contextes, en général, jamais seule car elle est couplée à des capteurs, des télécoms et du hardware.

Ce n'était bien entendu le seul objet de la semaine. Dans la tech en France, c'était aussi la semaine de la Blockchain avec des événements sur le sujet à s'y perdre, entre Bercy et Station F. Une multiplicité liée, d'après ce que l'on m'a raconté, à une sombre concurrence entre organisateurs d'événements et au souhait de l'Etat de rester plus ou moins neutre.

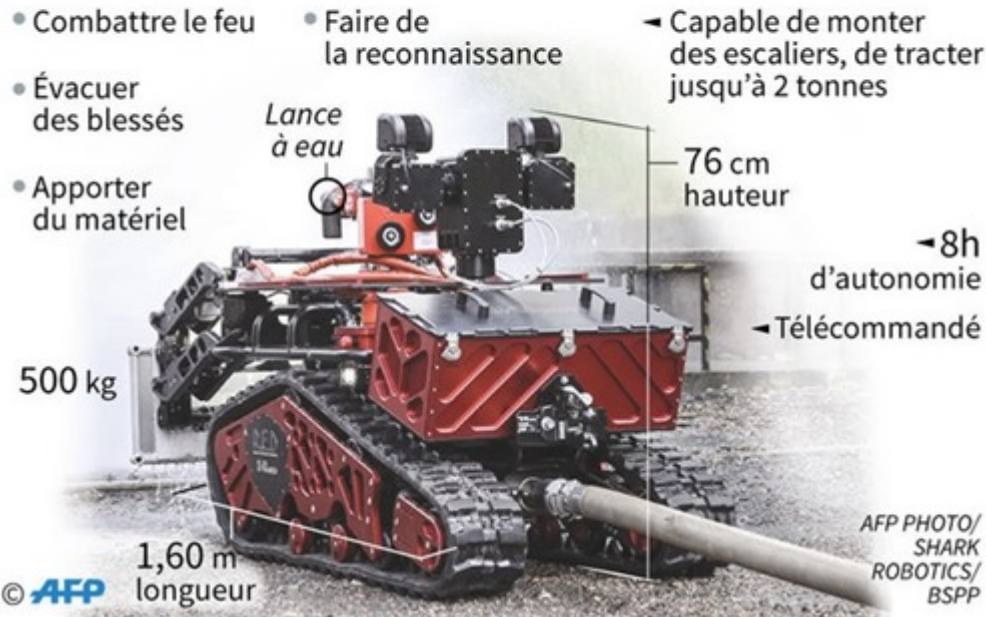
Lundi : Notre Dame

Le choc. L'incendie. L'IA était-elle au rendez-vous ? En apparence, seulement, dans un premier temps.

Le robot Colossus de la société **Shark Robotics** était utilisé par les pompiers pour éteindre l'incendie ou le contrôler à partir de l'intérieur de la cathédrale. En compagnie des pompiers, c'était avec ses 500 kg, l'un des héros de cette soirée et nuit tragique. Son géniteur est une TPE de 2 personnes créée en 2016 d'après **societe.com**. Basée près de La Rochelle, elle a déjà une belle brochette de clients français de l'Etat (DGA, Ministère de l'Intérieur) et d'entreprises du privé (énergie, industrie, ...). Le robot est-il pour autant intelligent ? Ses chenilles lui permettent de se déplacer librement et même, de gravir des escaliers. Mais il semble surtout que ce soit un engin télécommandé (vidéo). Il n'est pas entièrement autonome, même s'il est possible qu'il dispose d'une autonomie partielle pour éviter les obstacles et se déplacer.

Le robot Colossus

Utilisé par les pompiers dans l'incendie de Notre-Dame de Paris



Des drones ont également dû être utilisés par les pompiers pour inspecter la cathédrale pendant l'incendie. Ils s'appuyaient probablement sur des systèmes d'imagerie dans le visible et de thermographie. Mais l'interprétation des images était réalisée par des humains. Après l'incendie sont aussi mis en œuvre des systèmes de contrôle non destructif de la structure de la cathédrale. Mais pas forcément avec de l'IA.



L'IA interviendra probablement de différentes manières dans la reconstruction du toit. Un appel à projets est lancé. Les outils de la CAO (Conception Assistée par Ordinateur) seront bien entendu utilisés, à commencer par CATIA de Dassault Systèmes. Au minimum les créatifs pourront-ils s'appuyer sur des réseaux de neurones génératifs (GAN) pour créer des structures exploitant des bases de données de structures de cathédrales et autres monuments.

Au bout du compte, ce qui faisait la plus grande différence n'était pas l'IA mais l'ingéniosité et le courage humain. Pendant que le Président Américain utilisait sa bêtise naturelle pour proposer d'envoyer des Canadiens. Et les débats idéologiques sur le financement de la reconstruction du toit sont du même goût. L'IA n'y peut rien.

Mardi : j'avais quantique

C'était aussi une semaine quantique pour ce qui me concerne. Cela commençait lundi par une audition auprès de la mission parlementaire sur le quantique pilotée par la Députée **Paula Forteza**. J'ai y surtout parlé écosystème et rôle des différents acteurs, de la complémentarité nécessaire entre logiciel et matériel et aux besoins de formation rapide dans l'enseignement supérieur.

Le lendemain, et en compagnie de **Fanny Bouton**, nous avons formé une trentaine de femmes du numérique à l'informatique quantique. C'était l'une des actions lancées par l'association "**Quelques Femmes du Numérique !**" qui visent à renforcer la place des femmes dans les métiers du numérique, en particulier là où elles sont rares, comme dans l'entrepreneuriat et les métiers techniques. D'où l'idée de proposer cette formation d'une journée à l'informatique quantique qui est l'un des sujets les plus ardues qui soit dans le numérique. J'étais déjà intervenu de la sorte au sujet de l'IA en juillet 2018.



Nous étions accueillis par **Axelle Lemaire** dans les beaux locaux de Roland Berger à Paris. Les participantes étaient d'horizons très divers avec des *backgrounds* scientifiques ou pas selon les cas. Cette formation de 7 heures comprenait les interventions surprises de **Maud Vinet**, Directrice de Recherche du CEA-Leti, et de **Pascale Senellart**, chercheuse au CNRS et cofondatrice de la startup Quandela qui crée des sources de photons uniques. La première expliquait comment elle coordonnait la filière des qubits CMOS/spin d'électron avec plusieurs laboratoires à Grenoble. La seconde éclairait notre lanterne sur une autre branche des ordinateurs quantiques, ceux qui manipulent des photons.

Quel rapport avec l'IA ? Il existe, même s'il est un peu ténu. Les ordinateurs quantiques pourront un jour servir à accélérer certains types d'algorithmes d'entraînement de solutions de machine learning et de deep learning. Ce n'est pas forcément là où ils seront les plus utiles car peu de problèmes d'IA passeront du statut "impossible" à "réalisable" grâce au quantique alors que c'est le cas pour d'autres catégories de problèmes. Les calculateurs quantiques permettent notamment de simuler les lois de la physique pour évaluer la manière dont les molécules s'organisent. C'est utile en chimie inorganique et organique pour créer de nouveaux matériaux ou molécules. Le quantique permet de simuler le monde réel tandis que l'IA exploite souvent des données d'observation du monde réel pour les interpoler. Ce sont des approches complémentaires.



L'un des débats les plus riches et cryptique qui soit autour de ces concepts porte sur les différentes classes de complexité dans les problèmes à résoudre. En résumé, cela donne le schéma *ci-dessus* avec des problèmes qui sont résolus par l'Homme, l'IA ou l'informatique quantique, avec divers recouvrements. Les gains exponentiels théoriques que l'on pourra obtenir un jour avec des ordinateurs quantiques n'y feront rien : il subsistera des problèmes trop complexes pour eux et des problèmes dits indécidables. Conséquence de tout cela : l'Homme ne pourra jamais résoudre tous les problèmes même s'il fera sans cesse avancer la science ! La toute puissance de l'IA ou de quoi que ce soit d'autre est un mythe.

Si vous avez raté cette formation (où êtes un homme), vous pouvez vous rattraper en vous inscrivant à l'une des sessions de la même formation que je délivre avec Capgemini Institut les 27 mai, 23 septembre et 13 décembre 2019 ([programme et inscription](#)). Et là, ce n'est plus du *pro bono* comme mardi 16 avril !

Mercredi : éducation, culture et IA

Il y avait plus d'IA ce mercredi dans mes pérégrinations. Tout d'abord lors de la 7e **Journée de la Femme Digitale** organisée à la Maison de la Radio. Le terme "digital" m'hérissé le poil tellement il a une connotation "marcom" des usages du numérique mais soit. Il y avait de belles interventions pendant cette conférence comme celle de **Lubomira Rochet** (CDO de L'Oréal) sur la manière dont la confiance en soi se construit. Et ces 55 femmes entrepreneuses venues de tous les continents et rassemblée par **Total** (*ci-dessous*).



Il y avait aussi l'ultra-dynamique **Sophie Viger**, maintenant Directrice de l'école 42, qui expliquait comment elle a réussi à faire passer le pourcentage de femmes dans l'école de 15% à 26% en moins d'un an. Comme la mixité dans les métiers techniques du numérique n'est pas un problème simple, elle a avec ses équipes mis en place un panaché de recettes très pratiques. Les premières concernaient les conditions de l'accueil des femmes dans l'école elle-même aussi bien d'un point de vue pratique que du point de vue du comportement des uns et des autres. Les secondes portent sur les manières d'attirer les femmes dans les filières techniques du numérique.

Ce qui m'a le plus marqué et nous ramène à l'IA ? Cette intervention de Claude Terosier de **Magic Makers** avec ses cinq élèves allant de 7 à 15 ans. L'une d'entre elles, Tara, âgée de 15 ans, présentait sur scène son petit

véhicule autonome et expliquait comment elle lui avait appris à se mouvoir dans l'espace. Elle ne l'a pas dit sur scène mais elle développe tout cela en Python et en Tensorflow, le framework d'IA de Google. Elle connaît bien les enjeux de captation d'images et la gestion du moteur de règle de conduite. Respect ! (sa photo *ci-dessous* est de mon ami **François Tancré**)



Cela renforce un point que j'avais mis en avant dans **Le mythe de la fin du développement logiciel** en mars 2019 : on ne fait pas "du code", on crée des systèmes. Le métier de développeur associe à la fois celui d'architecte et de maçon du logiciel.

Au passage, si vous êtes chef d'établissement et souhaitez introduire l'offre de Magic Makers dans les activités extrascolaires des élèves, n'hésitez pas !



Le soir, je participais à un **débat** organisé par le **Centre des Monuments Nationaux** à l'Hôtel de Sully dans le cadre des Rencontres de l'Incubateur du Patrimoine, avec Marion Carré d'**AskMona** et Benoit Audigier du collectif **NoArtist**. Il était animé par **Fanny Bouton** et ce n'est pas un hasard ! Elle est aussi à la croisée des chemins entre IA et création. C'était l'occasion de réhumaniser les perceptions sur les notions de créativité autour de l'IA. Nous étions assez d'accords entre intervenants sur ce point. L'IA est une boîte à outil. La créativité impose au minimum la mise en place d'une boucle de feedback avec les humains. L'IA et notamment les réseaux génératifs utilisés dans l'écrit, la création graphique, musicale ou autre, ont toujours besoin d'être paramétrés par des développeurs et artistes et les créations qu'ils aident à engendrer sont ensuite évaluées et sélectionnées par les humains.

Comme le cinéma parlant, en couleur, la TV, les synthétiseurs en musique, les outils de création graphique (photo, dessin), l'imagerie de synthèse dans le cinéma ou les jeux vidéo, l'IA est une nouvelle palette d'outils qui élargit le périmètre de la créativité humaine.

L'IA génère des biais c'est bien connu (**exemple**). Comme elle exploite des données d'origine humaine, elle leur applique le principe du biais du rétroviseur qui lui fait reproduire des biais existants. On peut finalement expliquer les biais de l'IA avec la manière dont les biais humains se construisent eux-mêmes comme les biais de confirmation ou les biais de généralisation. L'analogie est même saisissante. Je vais la réutiliser à de nombreuses reprises dans mes futures interventions !

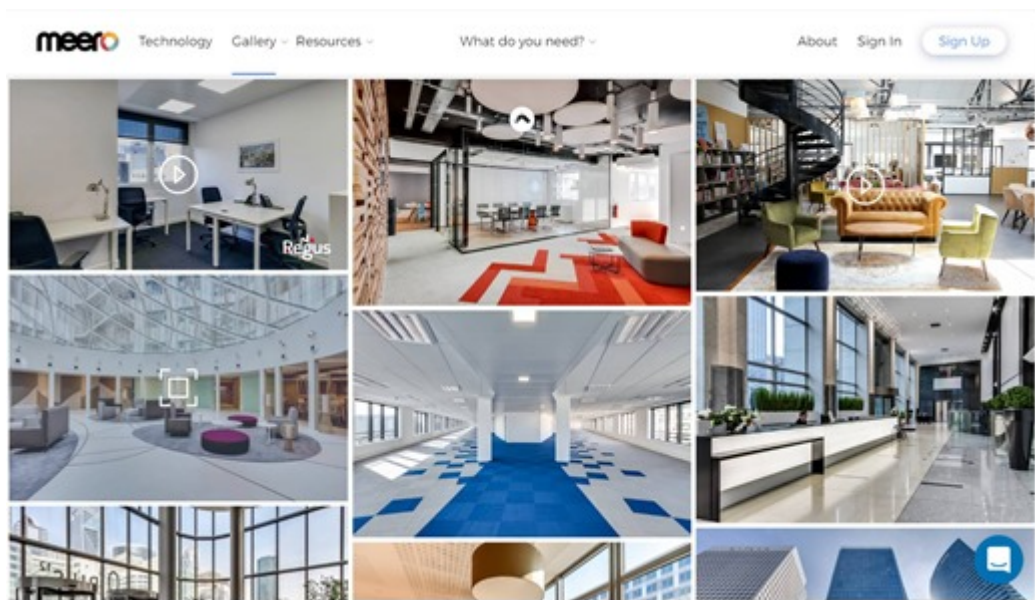
Jedi : AINight

Le matin, j'avais "5G" avec une conférence Orange Business au Palais Brongniart. Cela ne parlait pas trop d'IA, même si la 5G et l'IA font plutôt bon ménage dans nombre d'applications. Des annonces ? Pas vraiment. C'était surtout l'occasion pour Orange Business Services de mettre en avant les proof of concepts et projets de grandes entreprises clientes. La 5G est perçue comme apportant surtout de la valeur dans les applications b2b. Même s'il est probable que nombre de "killer apps" porteront rapidement sur les usages grand public.

L'après-midi et la soirée étaient eux consacrés à l'IA dans le cadre de l'événement **European AINight** organisé par Artefact pour le compte de France Digitale. La conférence introduite par **Cédric Villani** avait lieu en anglais car une partie de l'audience était non francophone, et notamment allemande. Le propos était européen mais toutefois dominé par des intervenants français, ne s'exprimant pas forcément bien dans la langue de Shakespeare. C'est très fatigant à écouter !

Certains cherchaient à valoriser de fumeux projets collaboratifs européens qui n'auront probablement pas un impact industriel majeur et à même de changer la donne vis à vis des géants américains ou chinois du secteur. J'ai toujours un arrière-goût désagréable dans ce genre de conférence. L'impression que les intervenants ont du mal à exprimer une vision ambitieuse. Ils exploitent plus l'état de l'art qu'ils n'ambitionnent de le faire avancer. Pourtant, le leadership européen devrait être incarné par des ambitions fortes, pas juste par un jeu de position.

Parmi les nombreuses interventions, il y avait celle de **Meero** et de son VP recherche en IA, Jean-François Goudou. Je les avais découverts car ils exposaient au CES 2019. C'est une startup parisienne qui veut révolutionner le workflow de production de photographies b2b et b2c en associant une plateforme d'intermédiation avec plusieurs dizaines de milliers de photographes dans le monde et un outil de traitement des photos à base d'IA. La startup a levé \$63,4M et elle est déjà implantée aux USA. C'est un bon exemple de startup qui exploite de l'IA mais dont l'IA n'est pas la seule partie de la proposition de valeur. Il y a aussi du business design et du process design dedans !



Pendant l'AINight étaient remis les "AI Awards" organisés par l'association Women in AI représentée par Moojan Asghari et Cag Gemini. Les trois lauréates gagnaient respectivement 5K€, 3K€ et 2K€. Il s'agissait de Camille Morvan de **Goshaba**, Priya Lakhani OBE de **Century Tech** et de Francesca Fedeli de **Mirrorable**. Ce sont toutes des entrepreneuses de startups exploitant de l'IA. **Goshaba** est une solution de matching de CV et de postes à pourvoir pour les recruteurs. Elle s'appuie sur plusieurs outils et méthodes : un outil de "smart screening", un jeu social mobile pour identifier des traits de personnalité de candidats et un outil de cartographie des talents dans un pipe de recrutement.

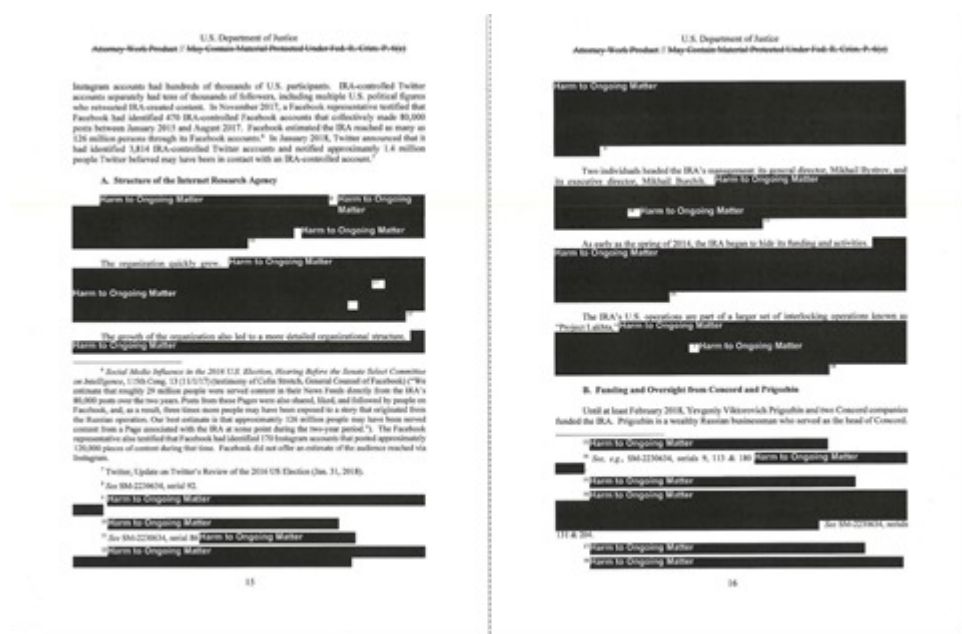
Vendredi : dépiater le Rapport Mueller

Publié jeudi 18 avril 2019 par le Ministère de la Justice américain, le Rapport Mueller était l'un des rapports les plus attendus pour déterminer ce qui s'était passé en 2016 et après autour de l'élection de Donald Trump en relation avec l'ingérence de la Russie. Le rapport était gris : il dévoilait les relations louches entre l'équipe de Trump et diverses parties Russes comme les tentatives d'obstruction de ce dernier pendant l'enquête. Mais au bout du compte, aucune inculpation. "No collusion, no obstruction".

En pratique, il y avait une coordination implicite entre les Russes et l'équipe de Trump mais pas suffisamment pour se faire prendre la main dans le sac : "While the investigation identified numerous links between individuals with ties to the Russian government and individuals associated with the Trump Campaign, the evidence was not sufficient to support criminal charges" (page 9). En gros, les personnages concernés n'ont pas été déclarés coupables, mais on n'est pas vraiment sûr qu'ils soient innocents. Le Rapport fait notamment état de conversations via des moyens de communication numériques cryptés qui n'ont pas pu être récupérées.

Quant à l'obstruction, elle est clairement démontrée dans le Rapport mais Mueller n'a pas jugé bon d'inculper le Président car les directives du DOJ ne le permettent pas. Le Président semble sorti d'affaire. Ou pas. Plus d'une quinzaine d'enquêtes se poursuivent, la majeure partie dont l'objet est encore confidentiel : "In the course of conducting that investigation, the Office periodically identified evidence of potential criminal activity that was outside the scope of the Special Counsel's authority established by the Acting Attorney General. After consultation with the Office of the Deputy Attorney General, the Office referred that evidence to appropriate law enforcement authorities, principally other components of the Department of Justice and to the FBI. Appendix D summarizes those referrals" (page 12).

Le DOJ publiait le rapport Mueller en PDF jeudi 18 avril vers 17h heure française. Il est officiellement téléchargeable [ici](#). Le document fait 448 pages mais il n'est pas "searchable". Les pages sont en fait fournies en mode "image". On sent la volonté de transparence ! Sans compter les 7,5% de texte noircis car ils concernent des enquêtes encore en cours, des procédés de renseignement inavouables ou des personnes qui ne sont pas inculpées.



Dans A Technical and Cultural Assessment of the Mueller Report PDF (avril 2019), la PDF Association a démolit le processus de création de ce rapport ! En effet, il résulte d'un scan à moyenne résolution (200 dpi) du document imprimé. Le fichier fait 142Mo. Par comparaison, le Guide des Startups 2019 de 550 pages fera moins de 30 Mo et le Rapport du CES 2019 de 416 pages faisait 43 Mo, le tout avec des illustrations en haute

définition.

Un rapport scanné pose plusieurs problèmes, notamment celui de ne pas être lisible par des non-voyants via des logiciels de speech to text. Il permet cependant d'éviter des bévues passées comme celle consistant à permettre que le texte en noir soit lisible avec un copier-coller ! La PDF Association rappelle qu'il existe des solutions techniques éprouvées pour produire des documents PDF de qualité, analysables avec des CONTROL-F / rechercher, tout en respectant la confidentialité des zones cachées.

De nombreux médias ont donc passé le rapport "format PDF image" au travers de logiciels de reconnaissance de caractères (OCR) pour le transformer en format PDF dans lequel des recherches sont possibles. C'est même devenu un véritable business, certaines éditions du rapport, même électroniques, étant payantes. C'est ce que raconte The Atlantic dans **The Irony of Mueller-Report Profiteering**. Le rapport est téléchargeable sur Scribd, mais en échange de vos coordonnées. Beurk. Pas beaucoup mieux chez Reason, *ci-dessous*. Enfin, j'ai trouvé une version correcte sans avoir à fournir de coordonnées, sur le site **Washingtonian** !



Le *data journalism* s'en est alors donné à cœur joie pour analyser le rapport. Il y a notamment le **New York Times** qui a dépiauté le texte dans tous les sens, l'a publié (sans PDF) mais en version analysable en ligne avec des commentaires explicatifs de journalistes ([lien](#)).

Axios a publié un outil permettant de voir d'un coup d'œil dans quelles pages des termes sont cités. Mais point de véritable IA à l'horizon.

C'est **Vox** qui est allé le plus loin avec beaucoup d'extraction de données du rapport pour analyser ce qui était barré, où et sur quel sujet. Ils ont utilisé une méthode un peu bourrine consistant à exploiter le scan des pages pour évaluer le pourcentage de pixels noirs en en déduire la quantité de textes barrés :



“Methodology: *We wrote a computer program to turn each page of the report into an image. Then we analyzed each pixel of the image to see whether it was black. Obviously, both the text and the redactions are black, which makes this exercise a bit hard. But we found that a page that is unredacted but full of text is about 3.5 percent black, and a page that is entirely redacted is about 34.5 percent black (because of border and imperfections in the redaction). So we were able to use these baselines to estimate how much each page was redacted. To determine the rationale for redactions, we used a similar method to count the number of times each rationale was used by counting pixel colors — and then double-checked it by going to each page by hand*

and recounting”.

En gros, on apprend dans leur explication que la méthode utilisée ne s’appuie pas du tout sur de l’IA mais sur des méthodes empiriques qui marchent tout aussi bien.



La meilleure analyse vient de l’avocat et essayiste **Seth Abramson**, considéré comme un conspirationniste paranoïaque par ses détracteurs, mais qui a vu juste sur un grand nombre de points jusqu’à présent et qui sont très bien décrits dans le Rapport Mueller. Son commentaire du Rapport est un thread record de 450 tweets publiés en une journée et il ne couvre que la partie “collusion” du Rapport. Vous pouvez le consulter grâce au site **Threader** qui compile des fils de discussion Twitter !

L’IA pourrait cependant jouer un rôle dans les phases suivantes d’exploitation du rapport et des autres documents produits par le *Special Counsel*. Des techniques à base de traitement du langage permettront peut-être de détecter des tournures de phrases, de comparer les déclarations sous serment avec les dires officiels de la Maison Blanche. Cette comparaison a déjà permis de démontrer les mensonges de Sarah H. Sanders, la porte-parole de la Maison Blanche. Dans la novlang de la Maison Blanche, “mentir” devient “slip of my tongue”. L’IA pourrait-elle détecter de tels mensonges éhontés ?

Un réseau de neurones génératif de type GAN (Generative Adversarial Network) pourrait-il remplir les parties noircies du rapport Mueller ? Probablement pas, mais cela pourrait être l’occasion de créer quelques parodies ironiques des capacités de l’IA !

D’autres pourront faire un décompte quantitatif du nombre de fois où le Président a dénoncé des “fake news” spécifiques qui ont été ensuite vérifiées. Malgré tout, si l’IA est un outil exploitable en pareille situation, l’intelligence naturelle est largement suffisante pour vérifier ou invalider tout cela.

Samedi : Gilets Jaunes

Et là, ce sera au tour des caméras de surveillance et aux logiciels associés de décompter les manifestants ! En coulisses, je ne serais pas étonné que le Ministère de l’Intérieur utilise des outils à base de traitement du langage pour analyser l’activité des gilets jaunes et black blocs dans les réseaux sociaux.

Sur ce, je retour au **Guide des Startups édition 2019** qu’il me reste à boucler avant de le publier pour la rentrée scolaire des franciliens, le 6 mai 2019.

Cet article a été publié le 20 avril 2019 et édité en PDF le 23 décembre 2021.
(cc) Olivier Ezratty – “Opinions Libres” – <https://www.oezratty.net>