By Alain Strowel, 8 May 2018

Quel menu pour nourrir l'intelligence artificielle? Pouvez-vous passer la carte?

Les outils d'intelligence artificielle (IA) se nourrissent de données. Mais le menu de données varie sensiblement d'une application algorithmique à l'autre. Et les recettes appliquées à ces données pour concocter des décisions varient aussi. Est-il possible d'obtenir une carte détaillée, avec une liste des ingrédients utilisés et une bonne explication des procédés?

Exemple 1: Les agents intelligents dans les jeux

La capacité de collecter, stocker et traiter d'énormes quantités de données est indispensable au développement d'outils intelligents. Lorsque Deep Blue (IBM) ou AlphaGo (DeepMind Google) gagne une partie d'échecs ou de go contre le meilleur joueur humain, c'est parce que l'agent intelligent a pu analyser, pour chaque mouvement des pièces, toutes les combinaisons et suites possibles. Et décider en conséquence du meilleur coup. Pour chaque déplacement sur le damier. Jusqu'à gagner. Le champ des possibles est certes vaste vu le grand nombre de déplacements de pièces et d'enchaînements potentiels, mais il est néanmoins étroitement délimité par le cadre du jeu.

Exemple 2: Les systèmes de conduite autonome

Dans la vie réelle, le plateau a une autre échelle. Sur les routes par exemple, les déplacements n'obéissent pas aux mêmes règles. Il y a certes un code de la route, mais il laisse place à des conduites très variées dans un environnement changeant. Finis les damiers, bonjour les nids de poule! Sans compter les piétons distraits, les chauffeurs empressés, l'angle mort du rétroviseur, la buée dans le pare-brise, … Les combinaisons ne s'enchaînent pas coup après coup, les déplacements des agents sont simultanés, les obstacles présents et imprévus, bref l'univers de la route est bien plus complexe et mouvant que celui d'un jeu de table.

Crédit photo: Usine nouvelle

Les algorithmes implémentés dans les objets intelligents que sont les voitures autonomes dictent directement des conduites pour les conformer aux règles de circulation. Le code numérique qui guide le pilotage de la voiture doit être univoque et précis afin d'être exécuté par la machine. Le code de la route avec ses instructions claires semble a priori respecter ces exigences. (N'est-il pas d'ailleurs souvent opposé, pour cette raison, aux vraies règles du droit? Qui se rappelle de l'avertissement d'un professeur de droit en première année: « apprendre le droit, c'est autre chose que d'étudier le code de la route! »).

Est-il pour autant évident de traduire les règles simples de roulage, telles que les limitations de vitesse, dans des instructions pour le système de contrôle et de pilotage de véhicules autonomes ? Une limitation résultant d'un pictogramme de limitation de vitesse peut être lue par un système de navigation, mais son encodage débouchant sur sa traduction dans des instructions à la voiture montre qu'il existe une multitude d'incertitudes que des programmateurs vont résoudre différemment : faut-il tolérer de modestes dépassements de vitesse ? Doit-on prévoir des phases de décélération progressive avant ou après le passage devant le signal de limitation ? Doit-on considérer qu'il y a une ou plusieurs infractions en cas de dépassement de vitesse pendant un certain laps de temps ? etc. (voir. L. A. Shay, W. Hartzog, J. Nelson et G. Conti, *Do robots dream of electric laws ? An experiment in the law as algorithm*, in R. Calo, M. Froomkin et I. Kerr (ed.), *Robot Law*, Edward Elgar, 2016, p. 274-305). Autrement dit, même la traduction d'un signal simple comme une limitation de vitesse peut donner lieu à des décisions de programmation variées et donc aussi à des parti pris. En outre, le système de navigation doit respecter beaucoup d'autres normes, en dehors du code de la route: des normes de sécurité (par ex. en cas de pluie ou de brouillard), voire des impératifs éthiques (par ex. préserver des vies humaines, plutôt que la "tôle"). L'automatisation de ces normes plus complexes peut révéler d'autres choix, et donc de biais, dans le chef des programmeurs.

Reste que les systèmes intelligents embarqués à bord des voitures autonomes doivent capter des masses de données afin d'en tirer les bonnes décisions de conduite.

newsroom.intel.com

La captation de données par les voitures intelligentes pose de multiples questions. Quel contrôle devons-nous avoir sur ces données? Et qui peut y avoir accès et à quelles fins: le constructeur pour affiner ses systèmes de pilotage et prévenir le hacking? le garagiste pour optimiser les entretiens? l'assureur pour adapter au mieux ses tarifs? le propriétaire du véhicule pour préserver sa vie privée? … Ces questions sont loin d'être résolues. Pour y répondre, il faudrait distinguer les types de données en jeu.

Exemple 3: Les algorithmes de sélection pour l'accès à l'enseignement

Dans d'autres cas de décision par algorithme, le problème semble plutôt se nicher dans les préférences sous-jacentes et les critères de sélection algorithmique. Ainsi en est-il de la controverse qu'a provoqué le système Admission Post Bac (APB) en France.

Ce système automatisé d'admission dans l'enseignement supérieur doit en principe "permettre à un maximum d'étudiants d'obtenir leur premier vœu. Mais lorsque les candidats sont trop nombreux pour la même formation (on parle de formation *"en tension"*), l'algorithme d'APB opère une sélection. Et ce, alors que le Code de l'éducation nationale garantit normalement à tout bachelier le droit à l'enseignement supérieur." (voir Franceinfo).

Lorsque la sélection a été effectuée par le système APB, le candidat n'a d'autre choix que d'accepter ou de décliner la seule préinscription retenue par l'algorithme. S'il accepte, il pourra s'inscrire dans la filière désignée. S'il refuse, il participera au *pool* suivant d'étudiants, quelques mois plus tard. Le portail APB a été développé car le système antérieur d'admission à l'enseignement supérieur avait été critiqué pour son manque de transparence et pour les risques de manipulation par le fonctionnaire procédant à la répartition entre les filières et universités. L'algorithme facilite une application uniforme des règles de droit, indépendamment des personnes impliquées. Il faut donc être prudent quand on oppose le gouvernement par la norme générale (par la loi) et la gouvernance individualisée par les algorithmes, le classement automatisé permet parfois d'assurer un traitement égalitaire et impartial qui peut faire défaut lorsque la norme générale est appliquée par l'administration.

Les futurs étudiants ont le droit de connaître les règles définissant le traitement algorithmique et ses principales caractéristiques de mise en œuvre en application du Code des relations entre le public et l'administration (J.M. Pastor, «Accès aux traitements algorithmiques utilisés par l'administration», *AJDA*, 2017, p. 604). Dans un avis de juin 2016, la CADA (Commission d'accès aux documents administratifs) a estimé que le code source de l'algorithme constitue un document administratif et doit être communiqué sur demande par la délivrance d'une copie sur support ou par courrier électronique. Ce code peut être utilisé à d'autres fins que celles des missions de service public mais dans le respect des droits de propriété intellectuelle que des tiers détiendraient sur ledit code source (*CADA, avis 20161989 du 23 juin 2016, Ministère de*

l'Education nationale). Cet avis de la CADA, ainsi que l'article 4 de la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique en matière d'ouverture des données publiques, ont été mis en œuvre par le décret n° 2017-330 du 14 mars 2017 relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique. Ce décret consacre, par le nouvel article R. 311-3-1-2 du Code des relations entre le public et l'administration, un droit d'accès de la personne sujette à la décision algorithmiquement fondée à diverses informations : le degré et le mode de contribution du traitement algorithmique à la prise de décision, les données traitées et leurs sources, les paramètres de traitement des données et leur pondération, ainsi que l'ensemble des opérations effectuées au cours du traitement. De nouvelles modalités de l'obligation de transparence administrative sont donc introduites au fur et à mesure du recours accru à des algorithmes guidant la décision administrative.

Ces obligations de transparence sont essentielles si l'on veut préserver l'autonomie des individus. Leur mise en oeuvre reste difficile. Il est en outre facile pour le législateur d'imposer la transparence en cas de décision administrative à l'aide d'algorithmes: on peut repartir des législations en matière d'accès aux documents administratifs. Que faire lorsque les algorithmes sont utilisés par des opérateurs privés?

Exemple 4: Les algorithmes de diffusion d'informations sur les plateformes

Depuis l'élection américaine de novembre 2016, on discute des fausses informations (les fake news), des biais dans la diffusion des informations en ligne et des risques que cela comporte pour le processus démocratique. La protection des données personnelles sur les grandes plateformes en ligne est aussi devenue une préoccupation majeure.

En avril 2018, Marck Zuckerberg, interrogé par les représentants au Congrès américain suite à l'affaire Cambridge Analytica, a mentionné l'IA (plus de 30 fois !) estimant que cette technologie « would one day be smart, sophisticated and eagle-eyed enough to fight against a vast variety of platform spoiling misbehaviour, including fake news, hate speech, discriminatory ads and terrorist propaganda » (*AI will solve Facebook's most vexing problems, Mark Zuckerberg says. Just don't ask when and how*, The Washington Post, 11 avril 2018). On ne peut s'empêcher de penser que s'en remettre à l'IA comme unique solution peut aboutir à diluer la responsabilité des grandes plateformes (cette foi en la technologie ne fait-elle pas penser à la croyance aveugle que le marché va régler tous les problèmes?).

Pour vérifier si les technologies d'IA peuvent résoudre les multiples dérives de l'information et des opinions en ligne, il faut à tout le moins que les plateformes de l'Internet qui servent de chambres d'écho, de relai de fausses nouvelles et de discours de haine, acceptent de rendre plus transparente la manière dont les algorithmes de propagation des contenus fonctionnent. Or elles refusent souvent de donner accès à ces informations aux chercheurs, invoquant notamment la protection de leurs secrets d'affaires, voire d'autres dispositions en matière de propriété intellectuelle (voir C. O'Neil, *Weapons of Math Destruction, How Big Data increases inequality and threatens democracy*, Crown, 2016, p. 29 et 185 ; R. Calo, *Artificial Intelligence Policy: A Primer and Roadmap* (August 8, 2017), disponible sur SSRN: https://ssrn.com/abstract=3015350). Dans d'autres cas, des opérateurs comme Google ont interdit à des chercheurs extérieurs de créer des faux profils afin de cartographier les biais du moteur de recherche (C. O'Neil, *op. cit.*, p. 211-212 ; voir aussi A. Strowel, *Quand Google défie le droit*, Larcier-De Boeck, 2011).

***

Depuis que ce billet a été mis en ligne (le 8 mai 2018), on a connu des développements intéressants sur l'exemple 3 (algorithme APB), voire quant à question envisagée sous l'exemple 4, à savoir la transparence des algorithmes de propagation de Facebook.

A. Deux points de mise à jour

- Pouvez-vous résumer les développements en France quant à la transparence du système automatisé d'admission dans l'enseignement supérieur?
- Est-ce que l'on a progressé en ce qui concerne la transparence des algorithmes de Facebook, Google et Twitter et de l'usage de l'IA comme remède à la propagation des propos haineux sur ces réseaux sociaux?

B. Questions ouvertes pour discussion

Une multitude d'outils intelligents sont en train d'être déployés dans divers secteurs. Les logiciels de jeu ne posent bien entendu pas les mêmes problèmes que les systèmes de conduite autonome. De même, l'usage d'algorithmes dans la décision administrative soulève d'autres questions de transparence que leur implémentation pour proposer des recommandations ou classer les informations en ligne sur les réseaux sociaux. A chaque fois, l'accès aux données ou aux recettes de fabrication des décisions algorithmiques sont centrales.

Voici trois questions à se poser:

- Y a-t-il un droit à l'explication des algorithmes dans les règles en matière de vie privée?
- Quelle est la pertinence des arguments juridiques invoqués par des opérateurs privés pour limiter l'accès aux algorithmes et/ou aux données?
- Faut-il revoir les exceptions aux secrets d'affaires et/ou à l'accès aux documents administratifs pour assurer la transparence des outils d'IA?

Merci pour la mise à jour et pour un début de réponse à ces questions ouvertes (Veuillez répondre en distinguant les aspects A. 1), 2) et B. 1), 2) et 3)).

***

Et enfin, une autre question touchant à la discrimination mais à laquelle vous ne devez pas répondre par écrit: de quel genre est l'IA? Plutôt du style costume-cravatte? Ou de l'autre genre?

iStockPhoto
source:http://www.adweek.com

source:
http://quillette.com/2017/12/14/irrational-ai-nxiety/

Peut-être que l'IA, souvent conçues par des développeurs masculins, rend les femmes invisibles? A considérer.

PS: des passages de ce post (sur le système APB) sont extraits de l'article co-écrit par E. Marique et A. Strowel, *Gouverner par la loi ou les algorithmes : de la norme générale de comportement au guidage rapproché des conduites, Dalloz IP/IT*, oct. 2017, n° 10, p. 517-521. (http://hdl.handle.net/2078.1/188689)