A N D S I

Association Nationale des Dirigeants en Sciences de l'Information
www.andsi.fr

Intelligence Artificielle (in)explicable & responsabilité… débats à l'ANDSI

Compte rendu rédigé par ANDSI & Pierre Delort

En bref…

Pierre DELORT, DSI et Professeur Invité à Télécom Paris, proposera un exposé autour des notions de data, d'Intelligence Artificielle explicable (XIA) et de responsabilité. Pour ce faire, il mobilisera de nombreux exemples issus de domaines distincts avant d'engager un débat sur la notion de responsabilité, fondamentale sur ce sujet d'XIA.

L'Association Nationale des Directeurs des Systèmes d'Information organise des débats et en diffuse des comptes-rendus, les idées restant de la seule responsabilité de leurs auteurs. Elle peut également diffuser les commentaires que suscitent ces documents.

Eléments d'introduction

En préambule, Pierre DELORT évoque l'utilisation du Big Data à travers l'exemple des travaux menés conjointement aux États-Unis par le CDC et Google dans le cadre du système Sentinel, dispositif qui rassemble 3 000 médecins. Sentinel assure un suivi de la propagation de l'épidémie de grippe saisonnière afin d'en prévoir le pic, en vue du déploiement de mécanismes en réponse : réservation de lits, vaccination, etc. Ensuite, le CDC collationne les données sur 9 zones géographiques et les produit avec un décalage de environ 10 jours avec le physique.

Afin de résorber ce décalage, un travail est notamment mené par Google et le CDC, en appui des données historiques du CDC, pour aller chercher, parmi les 50 millions de termes les plus recherchés sur Internet depuis 2004, les termes de meilleure corrélation (Pearson) avec l'historique de la grippe. A la pratique, il s'avère que les 45 premiers résultats ont tous un rapport avec la grippe, ce qui est rassurant sur la méthode suivie.

Depuis 2019, le modèle a été fortement amélioré notamment par l'utilisation de dossiers médicaux électroniques utilisés par 10 000 fournisseurs de santé aux États-Unis.

S'agissant de la thématique de la culture de la donnée en entreprise, une étude rapportée par *The Economist* a montré que le principal facteur limitant était de nature RH plutôt que technique. En outre, les différentes fonctions de l'Entreprise ne sont pas égales tandis que le partage est indispensable. Ce dernier donne la valeur des données.

Ainsi, après l'acquisition des droits politiques et économiques aux XIX^{e} et XX^{e} siècles, le XXI^{e} siècle sera peut-être marqué par l'émergence du droit à l'information. Un *Executive Order* du Président Obama prévoyait que, par défaut, les données devaient être disponibles et *Machine Readable*.

L'apparition du RGPD en Europe est liée à l'avènement de l'économie numérique. Les différents pays n'entretiennent pas le même rapport face à la protection de la donnée. L'Allemagne, en raison de son passé, est très attentive sur ce sujet tandis que la Chine l'est beaucoup moins. De surcroît, les menaces subies sur

l'intimité ne sont pas particulièrement récentes. Elles remontent à la fin du XIXème siècle et l'invention de la photographie et de la rotative de presse.

Bon nombre des modèles actuels tendent à accréditer l'idée selon laquelle l'anonymat des données personnelles est illusoire. En Europe et contre une possible tyrannie de l'État, des réflexions sont conduites à propos de l'anonymat et des données alors que de leur côté, les Américains contre cette même possible tyrannie de l'État, ont avancé le second amendement, avec pour conséquence l'armement massif des citoyens.

Et troisième exemple de l'influence de la culture sur ces sujets, la Cour suprême de l'Union Indienne a récemment reconnu l'intimité comme un droit constitutionnel, tout en avançant le droit des citoyens de l'Union Indienne aux progrès amenés par le numérique dont celui de l'accès aux connaissances.

L'Intelligence artificielle

Par ailleurs, Aristote et Platon ont mis en avant deux types de raisonnement des humains ; la déduction et l'induction. Ces deux éléments sont à l'œuvre dans le cadre de l'IA. Plus précisément, le *knowledge-based system* applique la déduction tandis que le *machine learning* applique l'induction.

Les premiers pas de l'IA (déductive) dans les années 1960-1990 ont été marqués par quelques innovations emblématiques, notamment un système mis au point par Schlumberger avec l'aide d'experts en IA du MIT afin d'aider la découverte de pétrole.

Le *machine learning* (IA inductive) fait quant à lui l'objet de développements marqués depuis les années 2010, dans un contexte d'accroissement des capacités de traitement, de transmission et de stockage de données. A ce jour, l'IA est très bien représentée par les réseaux de neurones profonds : le *deep learning*.

A titre d'exemple, un outil IA auquel est soumis un enregistrement audio très confus, où deux locuteurs s'expriment simultanément, dispose de la capacité d'identifier et séparer chacune des deux voix afin d'effectuer un filtrage et de les isoler, afin de les rendre intelligibles. Un tel résultat, présenté en conférence au collège de France, est remarquable.

La notion d'inintelligibilité

Pierre DELORT aborde la notion d'inintelligibilité d'un point de vue opérationnel. Un travail de comparaison de différentes technologies – régressions, système à base de règles, réseau de neurones – a été réalisé sur l'orientation de personnes souffrant de pneumonie à l'admission d'un hôpital. Il s'est avéré que les réseaux de neurones affichaient les meilleures performances globales maintenant ils renvoyaient systématiquement à leur domicile les personnes asthmatiques alors qu'une pneumonie est bien plus dangereuse dans ce cas. Les réseaux de neurones avaient été entraînés à partir de données selon lesquelles les personnes atteintes d'asthme étaient directement envoyées en soins intensifs, ce qui leur donnait un bon niveau de survie. En conséquence, la machine avait appris que l'asthme favorisait la survie des personnes souffrant de pneumonie, ce qui est… complétement faux et souvent mortel.

Un programme de recherche de la DARPA met en évidence l'existence d'une tension entre l'explicabilité et les performances. Par exemple, les arbres de décision affichent un très bon degré d'explicabilité, mais un assez faible niveau de performance, alors que c'est exactement l'inverse pour les réseaux de neurones profonds. L'explicabilité soulève ainsi plusieurs défis de recherche :

- comment produire des modèles plus explicables ?
- comment concevoir des interfaces d'explication ?
- comment comprendre les prérequis psychologiques pour des explications effectives ?

En outre, l'explicabilité de l'algorithme renvoie au fait que ce dernier fournit des moyens pour expliquer le résultat qu'il donne à chaque occurrence. Ensuite, se posent les sujets d'interprétabilité, consistant à savoir quoi faire de cette explicabilité et ceci repose notamment sur la transparence de l'algorithme.

Ouverture

En conclusion, Pierre DELORT propose d'ouvrir la réflexion sur le thème de la responsabilité devant l'inexplicable. Plus précisément, quelle responsabilité est-il possible de demander à un acteur qui prend des décisions sur la base d'algorithmes que personne ne comprend ? Yann LECUN établit un excellent parallèle avec les médicaments dont le fonctionnement est parfois peu compris. Ici, l'utilisation est validée par une AMM, incluant des démarches statistiquement valides. Yann LECUN explique qu'il pourrait être pertinent de réaliser le même genre d'étude pour les algorithmes que personne ne comprend. Le commentaire est que cela prend du temps.

Enfin, la responsabilité doit-elle incomber à l'individu décisionnaire ou à l'organisation (en tant que personne morale) à laquelle il appartient, comme les travaux de recherche de Pierre DELORT avec une banque l'ont avancé ?

Débat

Int : D'après l'Union Européenne, cette question trouve une réponse dans la loi de 1985 afférente au *opt-in*.
PD : Il incombait aux médecins de s'assurer qu'il n'y avait pas d'asthmatiques dans le cadre du test évoqué précédemment. Selon moi, cette notion de responsabilité est centrale. Dans l'exemple que j'ai évoqué à propos des asthmatiques, la responsabilité individuelle des médecins pouvait tout-à-fait être mise en cause dans la mesure où les asthmatiques auraient pu décéder, alors que le médecin était censé savoir leur surmortalité.

Int : L'exemple avancé à propos du médicament est intéressant. Lorsqu'un laboratoire met au point un nouveau médicament, ce dernier est soumis à toute une batterie de tests. A un moment donné, décision est prise de le mettre sur le marché, bien qu'il soit clair pour tout le monde que certaines personnes feront inévitablement une incompatibilité avec cette nouvelle molécule. Si nous l'acceptons avec les médicaments, pourquoi ne l'accepterions-nous pas avec l'IA ?
PD : En effet.
Int : Nous acceptons le risque. Toutefois, une famille n'acceptera pas le décès d'un de ses membres à la suite de la prise d'un médicament prescrit par un médecin.
Int : Pourtant, le médicament est autorisé par une AMM.
Int : Rien n'empêchera la famille dont il est question dans cet exemple de se retourner contre le médecin, car ce dernier aurait dû anticiper les potentiels effets indésirables qui étaient pourtant connus. Autrement dit, le médecin aurait dû anticiper le risque. Dans un autre registre, l'actuel débat sur la voiture autonome repose sur la notion de responsabilité. En cas d'accident, le conducteur d'un véhicule autonome sera responsable dans la mesure où celui-ci est présent à bord. D'ailleurs, des débats sont conduits à l'échelle européenne à propos de

l'interdiction de la voiture en conduite totalement autonome. La question se rapporte bien à la thématique de la responsabilité.

PD : Encore une fois, cette notion de responsabilité est centrale.
Int : Le véhicule autonome n'est pas plus dangereux qu'un conducteur qui conduit alors qu'il est sous l'emprise de l'alcool.
PD : Dans ce cas, la responsabilité est clairement identifiée. Il est nécessaire de savoir qui est responsable en cas de problème, cf. Interview d'Emmanuel MACRON donnée à Wired en 2018.
Int : En théorie, la voiture autonome sera moins vectrice d'accident que la conduite humaine. Il suffit qu'un accident de voiture autonome survienne pour que certains s'insurgent.

Int : Concernant les médicaments, la conduite d'études cliniques sur des patients Africains ou Chinois ne conduira pas à des résultats identiques. Le rôle du médecin est d'identifier les biais afin de correctement prescrire. Avec la voiture autonome, la situation est similaire. Les résultats seront différents selon que les véhicules roulent sur une autoroute ou dans un autre environnement.
Int : Je pense à un exemple récent où un laboratoire a décidé de modifier la formule d'un médicament permettant de traiter des problèmes de thyroïde. Ce laboratoire a réalisé de nombreux tests, et pourtant, il est apparu que la nouvelle formule entraînait des effets secondaires conséquents.
Int : Le panel d'individus concernés par les tests n'était pas suffisamment large.

Int : En cas d'accident avec un véhicule autonome, la responsabilité incombe-t-elle au concepteur au logiciel, au constructeur automobile ou au fournisseur de données ?
Int : Une directive de 1985 précise que le fabricant de la voiture est responsable.

Présentation des orateurs

Pierre DELORT est DSI, président de l'ANDSI et professeur invité à Telecom Paris – IP Paris. Ingénieur ETP, puis PhD de Mines Paris, il connut plusieurs positions de responsabilité en organisation d'entreprise, consulting puis Direction des Systèmes d'information ou Direction Scientifique.