

➤ n'apparaissant pas dans les algorithmes, on évite les difficultés liées à l'estimation de ce qu'il sait de sa cible, et la distinction entre quasi-identifiants et données sensibles, parfois peu évidente, n'est pas nécessaire. Mais cette simplicité est à double tranchant : des travaux publiés en 2011 ont montré que la non-prise en compte des connaissances annexes de l'attaquant et des relations potentielles entre les individus d'un jeu de données ouvre des failles dans la protection.

Un modèle d'assainissement universel des données reste donc chimérique. Chaque modèle a ses défenseurs et ses détracteurs, et est plus ou moins efficace selon la situation.

Outre le modèle d'assainissement, l'architecture de gestion des données a une importance. Les données nominatives sont souvent extraites du système informatique assurant leur usage quotidien (tel le serveur d'un centre de soin), puis copiées sous une forme pseudonymisée dans un entrepôt de données. C'est à partir de cet entrepôt que seront produits à la demande des jeux de données assainis pour différents destinataires. En France et en Angleterre, où le système de dossier médical personnel national fonctionne ainsi, les épidémiologistes peuvent par exemple recevoir des jeux de données plus précis que les industriels pharmaceutiques.

Toutefois, ce principe d'assainissement centralisé nécessite une fiabilité totale du gestionnaire de l'entrepôt de données. Or on constate régulièrement des fuites d'information sur de nombreux serveurs, dues à des négligences ou à des attaques. Même si les données stockées dans les entrepôts sont pseudonymisées, nous avons vu que cela n'apporte pas une protection suffisante. Il est donc légitime de s'interroger sur les risques de la centralisation.

L'ASSAINISSEMENT DISTRIBUÉ

Les statisticiens ont proposé un mécanisme d'assainissement décentralisé dès les années 1960, pour parer aux réticences à leur confier des données personnelles sensibles. Le principe est de perturber les données de chacun au moment de leur collecte, ce qui les protège avant tout enregistrement.

Cependant, on sait aujourd'hui qu'une perturbation indépendante de chaque donnée ne permet pas d'atteindre le niveau de qualité d'un assainissement réalisé sur le jeu de données dans son ensemble. La centralisation des données personnelles est-elle alors nécessaire à un assainissement de qualité ? Non, car il est aussi possible d'élaborer des mécanismes décentralisés où les perturbations ne sont pas indépendantes. En d'autres termes, la perturbation à appliquer n'est pas décidée localement, mais par une entité centrale. Celle-ci possède des informations sur toutes les réponses, sans connaître les réponses elles-mêmes.

Des algorithmes regroupés sous le nom de Secure Multi-Party Computation, qui font souvent intervenir des techniques de cryptographie, visent à assurer la confidentialité de l'assainissement distribué (où les calculs sont répartis entre plusieurs acteurs).

Les mécanismes distribués constituent un pas majeur vers la sécurisation du processus d'assainissement, mais leur mise en œuvre pose des problèmes de passage à l'échelle, car ils nécessitent des calculs importants et une forte connectivité des participants. Ils sont pour l'instant relégués au traitement de jeux de données peu volumineux.

Face à ce problème, notre équipe a développé une architecture de gestion de données personnelles fondée sur des dispositifs individuels, telles des clés USB, sécurisés. Nous avons montré que les calculs nécessaires aux algorithmes d'assainissement peuvent être répartis entre ces dispositifs et une entité centrale sans que celle-ci ne dispose des données personnelles non chiffrées. La puissance de calcul de l'entité centrale assure l'applicabilité à grande échelle et, grâce au fait qu'elle coordonne les dispositifs personnels, ceux-ci n'ont pas besoin d'être interconnectés en permanence. Cette architecture est capable d'appliquer des algorithmes d'assainissement à des jeux de données massifs, correspondant à plusieurs millions d'individus.

Une telle architecture décentralisée pourrait assurer la gestion des dossiers médicaux, des factures ou de tout autre type de dossier personnel. En pratique, les données seraient stockées localement sur les appareils de l'utilisateur et ne seraient jamais exportées sous une forme non perturbée ou non chiffrée. Aucun serveur ne les regrouperait toutes, mais les échanges entre l'entité centrale et les appareils des utilisateurs permettraient tout de même de collecter certaines informations, en respectant la vie privée.

Pendant la dernière décennie, une grande diversité de modèles et d'algorithmes d'assainissement de données ont été élaborés, afin de mieux prendre en compte les capacités de l'attaquant. Aujourd'hui, la pseudonymisation reste massivement utilisée, alors qu'elle ne garantit pas une protection suffisante ; dans le reste des cas, la *k*-anonymisation est le plus souvent appliquée, et les techniques plus complexes sont encore exotiques. Ces techniques doivent donc continuer à se répandre et à se perfectionner.

Pour autant, l'assainissement reste le meilleur compromis entre utilité et confidentialité des données, dont la protection absolue est illusoire. L'analyse des nouveaux gisements de données personnelles pouvant apporter des bénéfices notables, l'assainissement de données est une question sociétale avant d'être un défi scientifique. ■

BIBLIOGRAPHIE

T. ALLARD ET AL., METAP: Revisiting privacy-preserving data publishing using secure devices, *Distributed and Parallel Databases*, pp. 1-54, 2013 (à paraître).

B. C. CHEN ET AL., Privacy-preserving data publishing, *Found. and Trends in Databases*, vol. 2, n° 1-2, pp. 1-167, 2009.

A. GREENFIELD, *Everyware. La révolution de l'ubimédia*, Pearson, 2007.

C. DWORK, Differential privacy, *Proceedings of the 33rd international conference on Automata, Languages and Programming*, vol. 2, pp. 1-12, 2006.

L. SWEENEY, *k*-Anonymity: A model for protecting privacy, *Int. J. Uncertain, Fuzziness Knowl.-Based Syst.*, vol. 10(5), pp. 557-570, 2002.

L'Open Security Foundation, organisation non gouvernementale américaine, recense les cas les plus significatifs sur son site InternetDataLossDB.org

NOZHA BOUJEMAA



« Les algorithmes sont-ils transparents et éthiques ? Pour nous en assurer, nous avons besoin d'outils adaptés. »

Un algorithme peut-il être juste ?

Nozha Boujemaa : Pour répondre, un référentiel est indispensable. C'est précisément une des difficultés que l'on rencontre lorsqu'on s'intéresse à la qualité des algorithmes. La question à se poser est de savoir : « Pour qui ? ». La notion d'algorithme juste, ou loyal (d'une façon générale, on parlera de transparence), est différente selon que l'on se place du point de vue du consommateur du service adossé à l'algorithme ou du producteur de ce service.

Le consommateur individuel est un citoyen, et de fait les questions de transparence algorithmique ont souvent été soulevées par des affaires de libertés individuelles et de droits civiques. Quand l'utilisateur est une entreprise ou bien même les services de l'État, les problèmes possibles sont ceux, d'une part, de la concurrence déloyale, des abus de position dominante et, d'autre part, de souveraineté nationale.

BIO EXPRESS

17 JUILLET 1963
Naissance

2010-2015
directrice du centre
de recherche Inria
Saclay Ile-de-France.

2017
Directrice du projet
TransAlgo, de l'Inria
et de l'institut Dataia.

La question de la transparence est essentiellement celle de l'asymétrie informationnelle, entre celui qui propose un service logiciel et celui qui l'utilise, le second en sachant beaucoup moins que le premier. La réduction de cette asymétrie est cruciale : on doit savoir comment nos données sont calculées et par où elles transitent. Notons que la discrimination ou les biais potentiels ne sont pas toujours intentionnels.

Ces problématiques sont le cœur du projet *TransAlgo* que vous coordonnez. Comment est-il né ?

Nozha Boujemaa : Cette initiative fait suite à la loi pour une République numérique (promulguée le 7 octobre 2016) proposée par l'ancienne secrétaire d'État au numérique Axelle Lemaire, ainsi qu'au rapport commandé au Conseil général de l'économie (CGE) sur la gouvernance des algorithmes. La loi prévoit

notamment le droit à l'explication pour tout citoyen quant aux décisions algorithmiques prises par les services de l'État. C'est une avancée notable qui fait de la France un précurseur sur ces sujets aux niveaux européen et international.

En termes de protection des libertés individuelles, elle complète le Règlement général sur la protection des données (RGPD) adopté par le Parlement européen le 14 avril 2016. Ce dernier porte surtout sur les données personnelles, la portabilité, alors que la loi numérique s'intéresse à l'explication des décisions, celles-ci ne s'appuyant pas nécessairement sur des données personnelles. Le droit à l'explication est tout à fait original et anticipe bien l'évolution de la société où les algorithmes vont être de plus en plus amenés à prendre des décisions.

C'est une question de confiance dans l'État, et ses services, au moment où il entame sa transformation numérique. Il s'engage à cette occasion sur la transparence et l'explication des décisions algorithmiques. Il pourra alors répondre à des questions du type: «Pourquoi mon voisin et moi avons-nous été traités différemment alors que nous sommes dans la même situation?»

Très tôt dans la genèse de la loi numérique, la nécessité de l'accompagner par un volet scientifique s'est imposée. En d'autres termes, on a pris conscience des difficultés techniques qui allaient se dresser devant le législateur. De fait, l'explicabilité des algorithmes (expliquer le cheminement de la prise de décision et sa traçabilité) n'est pas une question résolue, tant s'en faut. De plus, une des recommandations phare, du CGE était la création d'une plateforme scientifique dédiée à la régulation et à la gouvernance des algorithmes. C'est là qu'entre en scène Inria, sollicité pour aider à la montée en compétence des services de l'État en ce domaine.

Mais il a d'abord fallu se mettre d'accord sur les objectifs ?

Nozha Boujemaa : De notre point de vue, il est préférable de ne pas parler de gouvernance ni de régulation des algorithmes, parce que c'est un peu utopique... Ce n'est pas le bon objectif à se donner. Je crois que nous avons été compris, même si chez nos interlocuteurs la tentation a été grande, je peux vous l'assurer, d'imaginer une autorité des algorithmes, un système de certification de transparence, avec des labels...

Ce n'est pas notre rôle en tant qu'organisme de recherche: chacun son métier et la régulation n'est pas le nôtre. En

revanche, nous sommes compétents pour répondre à des questions difficiles sur l'explicabilité, l'auditabilité et la loyauté des algorithmes, sur les biais possibles à la fois des données et des algorithmes, les deux étant liés dans ce qu'il vaudrait mieux appeler un système algorithmique. Cette dualité est essentielle. C'est sur ces questions que se penche *TransAlgo*.

Il s'agit donc d'ouvrir la boîte noire à laquelle on compare souvent les algorithmes ?

Nozha Boujemaa : Nous réfléchissons à des outils dits d'auditabilité pour essayer de comprendre le fonctionnement d'un algorithme. Pour ce faire, on n'a pas nécessairement besoin du code source. L'une des approches est la rétro-ingénierie, ou *reverse engineering*: avec des jeux de données tests bien calibrés et dont on maîtrise les variations, on étudie le comportement d'un algorithme afin d'en comprendre le mieux possible les rouages. Par exemple, on peut étudier un algorithme en ne faisant varier que les informations concernant la géolocalisation. On voit alors dans quelle mesure elle influe sur les résultats de l'algorithme.

La rétro-ingénierie est particulièrement utile pour un type de boîte noire, les algorithmes dont le fonctionnement est bien connu du concepteur, mais que celui-ci ne communique pas.

On doit bien maîtriser le protocole d'expérimentation et la méthodologie. Par exemple, avec un algorithme de classement, tel un moteur de recherche, les résultats peuvent varier simplement parce qu'au cours du temps, le corpus sur lequel opère l'algorithme a changé. Les précautions qui s'imposent parfois rendent indispensable le travail interdisciplinaire: des mathématiciens, des informaticiens, mais aussi des juristes, des sociologues, des économistes collaborent sur ces questions-là et encadrent ensemble les expérimentations.

« Le droit à l'explication anticipe l'évolution d'une société où les algorithmes vont prendre plus de place »

L'autre type de boîte noire concerne les algorithmes dont même le concepteur ne maîtrise pas tout le fonctionnement. C'est le cas de ceux issus de l'apprentissage profond, tels les réseaux de neurones, où les paramètres sont ajustés automatiquement, à mesure des exemples, pour remplir la tâche assignée. On doit alors se concentrer sur l'explicabilité. Des travaux de recherche sont indispensables pour rendre explicable ce qui ne l'est pas aujourd'hui.

Insistons sur le fait que tous les algorithmes de l'intelligence artificielle ne sont pas inexplicables, seuls ceux d'une famille particulière, relevant du *deep learning*, le sont.

Quelles approches sont explorées ?

Nozha Boujemaa : Dans cet élan, les approches d'*accountability* sont très intéressantes. Il s'agit de l'idée de rendre compte du comportement des algorithmes, d'évaluer leur loyauté en se penchant par exemple sur les données qui les ont alimentés et de repérer celles qui conduisent à des discriminations.

Il ne faut pas les éliminer, mais les identifier, les prendre en compte et maîtriser le processus de décision de sorte qu'elles n'influent pas dessus. On parle d'*Equity by design* («équité par construction»).

L'élaboration de nouvelles démarches d'apprentissage fait aussi partie de nos objectifs, afin que les questions éthiques et juridiques soient prises en compte dès le départ: c'est la *transparency by design*.

Les constructeurs d'algorithmes sont-ils prêts à vous suivre ?

Nozha Boujemaa : Au départ, plusieurs industriels n'ont pas très bien pris notre démarche. Pour eux, elle s'apparentait à un manque de confiance en leurs produits. Ils pensaient qu'on allait traquer les failles, les pointer du doigt. Mais nous ne sommes pas des gendarmes. Notre rôle est de mettre à disposition des outils logiciels au ➤

➤ service de tout le monde, citoyen, industriel, autorités de régulation...

Après un long travail de pédagogie, nous avons assisté à un changement des mentalités à notre égard. Les industriels ont compris qu'on leur proposait un moyen d'établir un lien de confiance avec le consommateur. Or la confiance incite à l'achat...

On peut faire le parallèle avec l'agro-alimentaire ou l'industrie pharmaceutique, deux domaines où l'importance de ce lien de confiance est reconnue depuis longtemps. D'ailleurs on m'a un jour fait remarquer que les médicaments nécessitent une autorisation de mise sur le marché, pourquoi n'en serait-il pas de même avec les algorithmes ?

Cette remarque est pertinente, car en effet, on met en circulation des algorithmes dont on sait peu de choses en fin de compte. C'est pour ça que l'État s'est posé la question de la régulation, et que nous avons proposé de travailler sur les outils de la transparence.

Pour certaines questions, des travaux de recherche sont encore nécessaires et peuvent réclamer du temps. Mais du chemin a été parcouru : des interlocuteurs hier méfiants ont compris l'intérêt de ces approches et proposent aujourd'hui de financer des thèses sur le sujet.

Comment faire admettre ce genre d'idée aux Gafam ?

Nozha Boujemaa : Il se trouve que ces Gafam (Google, Amazon, Facebook, Apple et Microsoft), et d'autres encore, sont à l'origine de Partnership on AI, une initiative visant à établir de bonnes pratiques dans les systèmes d'intelligence artificielle et à éduquer le public à ce sujet. Ils ont compris la vague citoyenne mondiale autour des questions d'éthique, de transparence, de discrimination, et ont pris les devants après la médiatisation de certains abus.

Il y a clairement une part de marketing, mais il y va aussi de leur survie commerciale. Aux États-Unis, de nombreux rapports et livres ont pointé du doigt les risques et certaines pratiques douteuses. Leur business menacé, ces grandes entreprises se devaient de réagir.

La clé de voûte de TransAlgo n'est-elle pas le volet pédagogique ?

Nozha Boujemaa : C'est en effet un point très important dans nos objectifs. Aujourd'hui, des groupes de travail thématiques ont commencé à réfléchir sur la définition des concepts que nous avons

évoqués (juste, loyal, éthique, responsable...) et à identifier ce qui est objectif, invariable. Cela pose la question de la relativité des concepts. Un site dédié est en cours de production, avec en page d'accueil cet aspect pédagogique essentiel.

Nous avons commencé à rédiger des textes synthétiques sur ces concepts et à sélectionner des cas d'usage emblématiques, par exemple sur la variabilité des prix, pour interpeller un large public. En complément, nous ouvrirons un espace de ressources avec des articles, des rapports, des vidéos, des logiciels en open source.

Et sur le plan scientifique ?

Nozha Boujemaa : En parallèle à ce volet pédagogique, nous lançons des groupes de travail scientifiques thématiques sur les algorithmes de classement, de recommandation, d'apprentissage, le contrôle de l'usage des données, les biais des données et des algorithmes. L'idée est de constituer une communauté scienti-

seront étudiées sans oublier leurs interfaces avec les sciences humaines et sociales. Une des priorités définies porte sur la transparence et l'éthique des systèmes algorithmiques. Le projet, prévu pour huit ans et demi et financé à hauteur de dix millions d'euros, rassemblera cent trente enseignants-chercheurs autour de quatorze partenaires académiques (Inria, CEA, École polytechnique, HEC, Ensae, Télécom Paris Tech, Centrale, Supélec, université Paris Sud, CNRS...).

Le numérique n'ayant pas de frontières, qu'en est-il de la dimension internationale ?

Nozha Boujemaa : Elle est fondamentale ! Nous ne pouvons pas nous contenter d'avoir notre propre définition de la transparence, de l'équité... Si c'était le cas, ce serait un coup d'épée dans l'eau. Aussi, dès le début du projet, nous avons discuté avec nos homologues du Data Science Institute, à l'université Columbia, aux États-Unis,

« Les Gafam ont compris la vague citoyenne mondiale portant sur les questions d'éthique, de transparence... »

fique de recherche sur le sujet réunissant plusieurs disciplines scientifiques (les télécommunications, la gestion de bases de données, l'apprentissage...), mais aussi des juristes, des économistes, des sociologues... Ce que produira cette plateforme scientifique sera ouvert à tous, vérifiable, testable.

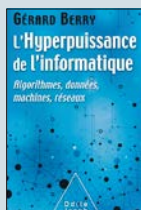
Plusieurs autorités de régulation sont aussi intéressées pour se joindre à cette plateforme et l'accompagner. Le Conseil national du numérique (CNNum) est également très impliqué. Le secrétariat d'État au numérique suit *TransAlgo* de près.

Mentionnons également Dataia, l'institut interdisciplinaire en sciences de données, intelligence et société que nous lançons (un des Instituts convergences financés par l'ANR et les Investissements d'avenir). Les sciences des données y

avec la National Science Foundation, aux États-Unis, et avec le programme Big Data Crest, de la Japan Science and Technology agency, au Japon, ainsi qu'avec l'institut Alan Turing, en Angleterre. L'idée de transparence algorithmique fait son chemin, malgré les différences marquées des sphères culturelles. Avec mes collègues américains, nous travaillons à ce que ce domaine bénéficie de crédits de recherche.

Nous avons aussi noué des contacts avec l'agence qui conseille le premier ministre italien sur l'agenda numérique, avec leurs homologues auprès du ministère du numérique espagnol... Il importe d'avancer ensemble. La mise en route est difficile, mais un mouvement international est indéniablement lancé. ■

PROPOS RECUEILLIS PAR LOÏC MANGIN



L'Hyperpuissance de l'informatique Algorithmes, données, machines...

GÉRARD BERRY

ODILE JACOB, 2017
(512 PAGES, 35 EUROS)

Depuis peu, on se rend compte à quel point l'informatique est en passe de transformer notre société, et même de la bouleverser. L'auteur, professeur au Collège de France, nous aide à nous repérer dans cette société où la science et la technologie informatiques mettent l'information au cœur de l'action. Pour ce faire, il passe en revue cinq domaines de transformations massives: les télécommunications, Internet, la photographie et la cartographie, la médecine. Pour finir, il donne sa vision de l'évolution de l'informatique.



Les Big Data à découvert MOKRANE BOUZEGHOUB ET RÉMY MOSSERI (DIR.)

CNRS ÉDITIONS, 2017
(368 PAGES, 39 EUROS)

Depuis les tablettes mésopotamiennes jusqu'à l'intelligence artificielle, l'humanité n'a eu de cesse de collecter des données. Mais le phénomène s'est amplifié depuis l'essor de l'informatique et d'Internet qui nous a conduits vers l'ère des *big data*. Dans cet ouvrage, un large panel d'experts présente, à travers 150 articles synthétiques, l'état de l'art sur cette révolution: les défis soulevés, les bienfaits attendus, les dérives à éviter, les risques à maîtriser... Un livre indispensable pour saisir tous les enjeux et participer aux débats de société engagés.



Le Temps des algorithmes GILLES DOWEK ET SERGE ABITEBOUL

LE POMMIER, 2017
(192 PAGES, 17 EUROS)

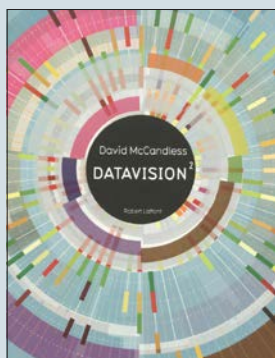
Les algorithmes ont envahi le débat public et les discussions. C'est sans doute parce qu'ils inquiètent autant qu'ils fascinent. Nous nous réjouissons qu'ils nous facilitent la vie, mais redoutons qu'ils nous asservissent... Les auteurs, spécialistes reconnus, nous invitent à sortir de cette vision manichéenne et proposent un nouveau regard sur ce qui, après tout, n'est que créations de l'esprit humain: les algorithmes sont ce que nous avons voulu qu'ils soient. L'ouvrage nous enjoint à ne plus les subir, mais à chercher à les comprendre. C'est ainsi que nous pourrions être maîtres de notre destinée.



Les Statistiques en BD LARRY GONICK

LAROUSSE, 2016
(256 PAGES, 17,99 EUROS)

L'auteur, dessinateur, professeur et mathématicien fait le pari de vous faire aimer les statistiques. Avec un humour décalé, il dédramatise l'univers complexe de cette branche des mathématiques aujourd'hui omniprésente: économie, finance, technologie, informatique, biologie, communication. Nous sommes entourés de statistiques, alors autant ne pas se laisser piéger par leurs chiffres et acquérir les grands principes de base pour les... déchiffrer, surtout si c'est en s'amusant! Les lois de Bernoulli s'offrent à vous, ne résistez pas.



BEAU LIVRE

Datavision². Le savoir est un art DAVID MCCANDLESS

ROBERT LAFFONT, 2014
(224 PAGES, 23,50 EUROS)

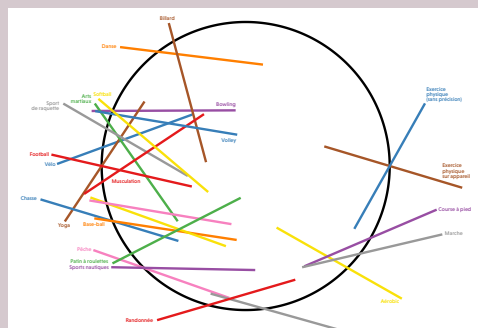
L'auteur, designer et pionnier de la datavisualisation, propose ici ses plus belles infographies. Elles démontrent que les images offrent une approche plus directe, mieux hiérarchisée, plus facilement mémorisable et plus séduisante du savoir universel. L'ouvrage offre une réponse graphique et esthétique à toutes sortes de questions. Au hasard: une guerre nucléaire totale est-elle envisageable? L'huile empêche-t-elle les pâtes de coller? Quels sont les mots de passe les plus utilisés sur le Web? Quelles villes seront sous l'eau en 2100?

RENDEZ-VOUS

P. 114

DONNÉES À VOIR

DES INFORMATIONS
SE COMPRENNENT MIEUX
LORSQU'ELLES SONT MISES EN IMAGES



P. 118

SPÉCIMEN

UN ANIMAL ÉTONNANT CHOISI
PARMI CEUX PRÉSENTÉS SUR
LE BLOG « BEST OF BESTIOLES »



P. 110

REBONDISSEMENTS

DES ACTUALITÉS SUR
DES SUJETS ABORDÉS
DANS LES HORS-SÉRIES PRÉCÉDENTS



P. 116

LES INCONTOURNABLES

DES LIVRES, DES EXPOSITIONS,
DES SITES INTERNET, DES VIDÉOS,
DES PODCASTS... À NE PAS MANQUER



P. 120

ART & SCIENCE

COMMENT UN ŒIL SCIENTIFIQUE
OFFRE UN ÉCLAIRAGE INÉDIT
SUR UNE ŒUVRE D'ART

