

## Compte-rendu d'ouvrage

Référence : Cathy O'Neil, *Algorithmes, la bombe à retardement*. Paris : Les Arènes, 2018. Préface de Cédric Villani.

Titre original : *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*.

### L'auteure

L'auteure connaît bien le sujet des algorithmes. Mathématicienne de formation, elle a travaillé comme analyste dans un fonds spéculatif (hedge fund), milité avec le mouvement « Occupy Wall Street), travaillé comme experte en données au sein d'une start-up publicitaire et participé à la formation de journalistes.

### L'ouvrage

L'ouvrage se déroule en dix chapitre chacun se référant à un domaine d'activités (finance, formation, commerce, etc.) et présentant moult exemples.

Etonnamment les algorithmes n'occupent pas régulièrement le devant de la scène<sup>1</sup>. Le cadre général est plutôt une analyse féroce de l'organisation inégalitaire de la société, américaine avant tout, dont certains usages des algorithmes constituent la phase ultime. La dédicace en est témoin : « A tous ceux que l'on donne comme perdants ».

Venons-en à la part des algorithmes.

Ceux-ci régissent la vie de la population dans un nombre important de domaines. Accès au crédit, à l'emploi (dépouillement automatique de CV), à la formation, etc. La sécurité et le système judiciaire en font largement usage. Presque toutes les activités et institutions, dont les universités, sont soumises à des évaluations reposant sur des systèmes automatiques. Les e-scores pullulent.

Parmi les algorithmes, l'auteure distingue les ADM (Armes de Destruction Mathématique) dont la définition se résume en trois éléments plus ou moins conjugués : opacité, échelle, nocivité.

Avant de passer en revue ces qualificatifs, il peut être utile de présenter un système qui n'est pas catégorisé comme ADM. C'est celui qui permet de prévoir des résultats à des match de base-ball. Il se base sur des statistiques connues des performances des joueurs selon leur position dans le jeu. Avant l'automatisation, cela permettait déjà certaines prédictions à l'instinct. Ce système n'est pas opaque et n'est pas nocif dans la mesure où ses performances peuvent être jugées dans la réalité qu'il modélise (les compétitions de base-ball) et être amélioré en conséquence.

L'opacité<sup>2</sup> est liée au fait que les traitements ne sont pas rendus publics et que les données utilisées ne sont pas non plus disponibles. Très souvent, des données de substitution sont utilisées. Par exemple une note de e-crédit peut-être utilisée comme un indice de moralité. Même de bons algorithmes raisonnables peuvent fonctionner avec des données qui reposent sur des choix d'humains faillibles. C'est d'ailleurs pourquoi les systèmes automatiques tendent à reproduire des jugements a priori. Un habitant d'un quartier défavorisé risque d'être a priori être mauvais payeur, récidiviste, etc.

---

<sup>1</sup> Par exemple, l'auteure explique comment les fonds spéculatifs sont alimentés par les retraites de petites gens et sont traités par Wall Street comme de la Dumb money, l'argent stupide des investisseurs inexpérimentés.

<sup>2</sup> A noter que dans le monde des informaticiens un logiciel qui cache les traitements est souvent qualifié, au contraire, de transparent.

Par ailleurs, des erreurs de données peuvent entacher le fonctionnement du système : garbage-in garbage-out (ce qui entre faux ressort faux) comme disent les gens du métier.

L'échelle est liée au fait que l'usage de l'algorithme peut petit à petit se substituer à l'examen des lois. C'est le cas du système LSI-R qui juge de la probabilité de récidivisme. Il est plus économique à utiliser que la convocation d'un groupe d'experts mais il rend des jugements, basés sur de données pas forcément fiables, que l'on ne peut pas remettre en cause (l'ordinateur est objectif).

Pour illustrer la notion de nocivité, le cas du logiciel IMPACT est présenté. Il est utilisé dans le district de Washington D.C. pour évaluer les enseignants et permet de renvoyer quelques pourcents de ceux jugés les moins performants. Il est basé sur le modèle de la valeur ajoutée. Il compare les résultats des élèves en fin d'année par rapport à ceux de l'année précédente. Son principal défaut est que les résultats sont calculés sur 20 à 30 élèves ; il peut donc y avoir des volées particulières. Par ailleurs, des enseignants pourront être tentés de gonfler les résultats en fin d'année ce qui va préteriter l'enseignant du degré suivant. De plus, les mauvais enseignants étant renvoyés, il n'y a pas moyen de vérifier si le système s'est fourvoyé. Il n'y a pas de remontée de l'information. Il est la référence et s'auto-justifie. De plus, le traitement est totalement opaque. Aux enseignants qui se renseignent à ce sujet, on répond que c'est un algorithme très complexe. Il est ainsi difficile de mettre en doute un score obtenu.

Un autre logiciel de ce type est Cataphora qui repère dans une entreprise les bons employés générateurs d'idées. PredPol en est un autre qui traite de la prévention et répartit les forces de polices en conséquence en se concentrant surtout sur les quartiers pauvres.

Le système de classement des universités est aussi un ADM selon l'auteure<sup>3</sup>.

Lorsque le classement des universités s'est avéré nocif, l'administration d'Obama n'a pas réussi à élaborer une nouvelle méthode de classement. Les présidents des universités s'y sont opposés eux qui avaient passé des dizaines d'années à réaliser des corrections. Le résultat a été ce que l'auteur appelle un anti ADM, c'est-à-dire la mise à disposition de masses de données qui permettent aux étudiants de poser leur propres questions.

Certains ADM ont bien un feed-back qui permet de juger de son fonctionnement. Mais il n'est pas lié aux hypothèses de base. Dans le domaine de la finance, par exemple, l'efficacité du système est jugé en fonction des gains qu'il permet de réaliser. Les faux positifs ou négatifs par rapport aux prévisions ne sont pas en compte dans l'évaluation. Par exemple le nombre de sujets jugés solvables alors qu'ils se révèlent en défaut de paiement n'est pas pris en compte. Il en va de même pour les sujets jugés insolubles alors qu'ils se révèlent être bon payeurs. La crise de 2008 est un bon témoin des résultats de cette pratique.

Les systèmes nocifs ont tendance à renforcer les pré-jugés qui leur ont servi de modèle de base<sup>4</sup>. Ils ont tendance à accroître les inégalités selon le schéma : les milieux populaires sont traités par des systèmes automatiques (accès aux crédits, aux places d'université, à l'emploi) pleins de préjugés alors que les classes plus favorisée peuvent bénéficier de traitements plus nuancés menés par des humains.

Les ADM transposent sur les gens les traitements faits aux objets, par exemple la technique du juste-à-temps qui organise le temps de travail du personnel à horaire variable.

Un chapitre est consacré au mini-ciblage. Bien connu dans le domaine de la publicité, où son usage peut être justifié, par exemple en proposant des bons de réduction aux acheteurs les plus susceptibles de changer de marque mais pas aux habitués. Ils peuvent être abusif lorsqu'ils

---

<sup>3</sup> Le groupe « University Apollo » maison mère de l'université de Phoenix a dépensé en 2010 2225 US\$ par étudiant en marketing et 892 US\$ pour l'enseignement.

<sup>4</sup> Le secteur du Big Data exploite les mêmes talents issus des universités formant les élites que le domaine de la finance. L'auteure ne semble pas apprécier leur arrogance.

« s'attaquent » aux personnes vulnérables (comme le marketing téléphonique auprès des vieilles personnes).

Cette technique s'utilise en politique. En juillet 2011, l'équipe d'Obama recrutait via LinkedIn des experts en statistique, en machine learning, en extraction de données, en analyse textuelle et en analyse prédictive pour contribuer à la stratégie électorale, avec comme mission de cibler les indécis. Les opposants organisaient de leur côté des *fake-news* micro-ciblées informant qu'Obama était musulman et né en dehors des Etats-unis.

L'équipe d'Hillary Clinton s'est appuyée sur cette méthodologie en engageant une start-up spécialisée dans le microciblage. L'équipe de Ted Cruz, un concurrent républicain, semble avoir eu recours aux service de la société Cambridge analytica. L'auteur de fait pas mention de la campagne de Donald Trump<sup>5</sup>.

Le micro-ciblage peut aller jusqu'à proposer des pages web différentes selon l'utilisateurs.

Un chapitre est consacré à Facebook et c'est très intéressant. Les détails apportés sur le fonctionnement du système dépassent largement les échos diffusés par les médias lors du « vol » de données et la manipulation, présumée, par la société Cambridge Analytica.

Contrairement aux enquêtes psycho-sociologiques menées d'habitude sur quelques dizaines de sujets, les laboratoires de Facebook ont à disposition quelques dizaines de millions de sujets. Google utilise également les expériences A/B (une partie des utilisateurs reçoivent une version A d'un document et une autre partie une version B) pour juger de l'efficacité d'une présentation, d'une mise en page, etc.

Pour faire des expériences, les utilisateurs américains ont la possibilité de cliquer sur un bouton « j'ai voté », ainsi parmi d'autres expériences, les chercheurs engagés par la multinationale peuvent contrôler, par exemple, quel est l'effet de savoir que des amis ont voté.

Un outil « mégaphone électoral » a été expérimenté dès 2010 (législatives), 2012 (élection présidentielle).

De façon générale, il est possible de manipuler les informations qui apparaissent chez chacun. Ainsi, il est possible de choisir les utilisateurs qui vont recevoir une pétition, une publicité, etc. et ainsi de juger la façon dont les utilisateurs sont influencés. De même, il est possible de déterminer comment « l'amitié » est un facteur qui influence l'opinion. Ces informations sont à la base du micro-ciblage.

Sachant que la société est cotée en bourse, le feedback utilisé pour « améliorer » le système est évidemment les résultats financiers et non la qualité sociale du service, même si cette dernière peut jouer un certain rôle. Il s'agit pour Facebook de jongler entre réglementation, réputation, utilité et gains.

Toutes les entreprises des GAFAM et d'autres, sont également concernées par ces pratiques.

### **Moralité**

Si le tableau concerne les Etats-unis, il nous touche également par toutes les entreprises qui nous manipulent depuis la Californie (nous sommes tous des californiens). Uber, Airbnb, Tripadvisor, etc. sont largement utilisées. Ce sont en général des services jugés très utiles. Il serait très intéressant de s'interroger sur leur modèle et leur impact global. Il en va de même d'autres systèmes, plus spécialisés tel que Researchgate dans le domaine de la recherche.

Il y a certainement des systèmes propres à nos sociétés. Derrière les appels téléphoniques peuvent se trouver des systèmes de profilage qui déterminent les personnes susceptibles d'être

---

<sup>5</sup> Le livre a été publié dans sa version originale en 2016 à l'époque des élections. Une suite serait la bienvenue pour éclairer avec la même richesse de détail ce qui concerne la campagne de D. Trump

intéressées ou être convaincues (les plus fragiles!). Le prévention en matière criminelle est aussi à l'ordre du jour<sup>6</sup>.

Pour continuer le livre de Judith Duportail : L'amour sous algorithme (Ed. La Goutte d'or) analyse le système Tinder et dénonce le modèle sous-jacent et l'opacité des algorithmes qui délivre des e-scores de désirabilité. (DLI)

---

© 2019, SENS & l'auteur

---

<sup>6</sup> <https://www.rts.ch/info/regions/autres-cantons/9236453-la-prediction-des-crimes-via-logiciel-ne-seduit-pas-en-suisse-romande.html>