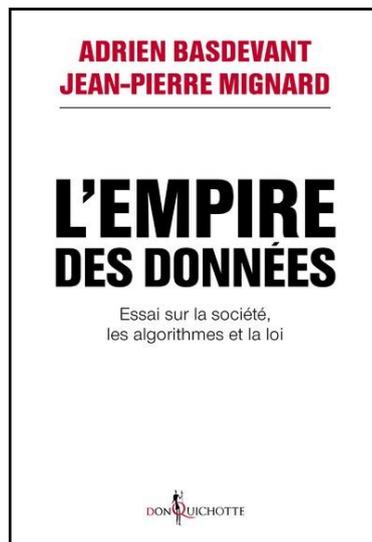


## NOUS AVONS LU

« *L'empire des données – Essai sur la société, les algorithmes et la loi* »,

Adrien Basdevant et Jean-Pierre Mignard, Don Quichotte éditions, mars 2018



Deux avocats du Barreau de Paris ont rédigé ce livre brillant, très détaillé, facile à lire, mais inquiétant.

Depuis John Graunt, au XVII<sup>ème</sup> siècle, le petit mercier à l'origine du retour d'expérience, les méthodes statistiques sont de plus en plus utilisées et prennent de plus en plus d'importance. Avec la transformation numérique, parce que les coûts de stockage des données sont bas, parce que les outils de collecte et les logiciels d'analyse sont maintenant disponibles et performants, on assiste à une croissance *runaway* exponentielle des volumes de données numériques, que l'on désigne par *big data*. Les applications sont multiples. On en trouve dans chaque secteur d'activité, par exemple : la main-

tenance préventive et la maintenance prévisionnelle, l'exécution des opérations boursières, les *clusters* de symptômes en médecine, la prévention des catastrophes naturelles et des catastrophes industrielles,...

Les auteurs nous emmènent de la collecte aux modèles « boîtes noires », efficaces, sans pareil, quand il s'agit d'anticiper des situations redoutées. Les modèles statistiques piègent les situations fâcheuses, nous dictent la conduite, nous révèlent des comportements (techniques comme humains), permettent de mieux comprendre les complexités, les effets, mais semblent oublier les causes (directes ou profondes) et la compréhension physique des événements. Serait-ce maintenant superflu ?

Ces modèles permettent de définir une moyenne, une classe moyenne, un « homme moyen », une entité normale au sens gaussien. Quid des situations extrêmes ? « On voit poindre le passage de la statistique descriptive à la statistique prescriptive » (page 37). « Le *big data* porte ... la croyance de pouvoir accéder directement au réel, en collectant et en mesurant tous les signaux, sans plus devoir interpréter ni questionner leurs contenus. On s'en remet à eux de manière quasi-mystique... En réalité, la véritable avancée de cette approche est de donner accès à des phénomènes auparavant non perceptibles » (pages 71-72).

Les auteurs posent le problème de la propriété des données. De notre point de vue les données sont la propriété de celui qui les collecte. Celui qui possède les données, qui a aussi accès aux structures open data, possède la gouvernance par les nombres. « La puissance du coup data réside dans la possibilité de tirer de l'information de données autrefois considérées comme insignifiantes. Des données qui étaient auparavant anodines sont désormais potentiellement une source d'information sur nos identités multiples... » (page 115).

La fin de l'ouvrage est consacrée aux aspects juridiques, politiques ou éthiques (le *coup data*, la fin d'une certaine conception de la démocratie, quand l'algorithme se substitue à la loi,... (page 247)). Le *coup data* est-il un monde sans hasard ? Le hasard et l'incertain sont vertueux, ils ont permis des progrès importants (c'est la sérendipité (*serendipity*)). Il faut défendre l'individuel face à une société de surveillance. L'utilité des méthodes statistiques, le recours à des méthodes de prévention ne sont absolument pas en cause. La compréhension physique, la compréhension des causes ne sont pas un domaine superflu. Les algorithmes du *big data* ne supprimeront pas le hasard. Dans notre discipline où le retour d'expérience (les données d'entrée), qu'il soit technique ou humain, est notre matériau de base, la transformation numérique va bouleverser nos approches. Nos deux groupes de travail et de réflexion ont un avenir radieux.