

LE BIG DATA VA-T-IL CHANGER LA SCIENCE ?

PAR ÉTIENNE KLEIN
PHYSICIEN



Tout le monde en parle, et pour cause : le big data, couplé à l'intelligence artificielle, a un impact majeur sur les technologies, le travail, les emplois, l'économie, la médecine, les relations interpersonnelles, et aussi, à l'évidence, sur les sciences.

Depuis Galilée, les deux grands piliers de la science moderne ont été la théorisation et l'expérimentation : on élabore des hypothèses conduisant à la formulation de lois, qui sont ensuite testées grâce à des expériences spécifiques, souvent très éloignées de la simple observation. Par exemple, lorsque,

en 1604, Galilée énonce que tous les corps, quelle que soit leur masse, tombent à la même vitesse, il formule une loi qui n'est ni conforme aux données empiriques, ni directement observable (elle ne peut l'être que dans le vide) : elle ne sera expérimentalement vérifiée que plus tard.

Avec l'arrivée du big data, ce modèle demeure-t-il toujours vrai ? En juin 2008, Chris Anderson, rédacteur en chef de *Wired Magazine*, publiait un article au titre provocateur, qui semblait répondre que non : « The End of Theory : The ►►

Le big data. « Un ensemble énorme de données brutes et silencieuses qui mémorisent les « traces » de l'activité d'êtres vivants, de machines, d'objets, et leurs états successifs. » ►



Data Deluge Makes the Scientific Method Obsolete». Il y défendait l'idée que lorsque nous disposerons de suffisamment de données, les nombres parleront d'eux-mêmes et les corrélations qu'ils dévoileront remplaceront les relations de causalité que manifestent les lois théoriques. La science changerait alors de visage puisqu'elle pourrait se dévelop-

« **CE N'EST PAS PARCE QU'IL Y A DES
GRENOUILLES APRÈS LA PLUIE QU'ON A LE
DROIT DE DIRE QU'IL A PLU DES GRENOUILLES.** »

per sans faire d'hypothèses et sans s'appuyer sur des théories explicites.

Mais qu'est-ce que le big data ?

Un ensemble énorme de données brutes et silencieuses qui mémorisent les « traces » de l'activité d'êtres vivants, de machines, d'objets, et leurs états successifs. Leur collecte est principalement automatique, de sorte que le traçage se fait sans acte volontaire (s'il s'agit d'un être vivant) de la part de l'entité tracée. Ces données sont analysées par des algorithmes pour y détecter des régularités, par exemple dans le comportement des consommateurs, des machines, des indices économiques, du trafic routier... À partir de ces régularités, ils infèrent des règles prédictives que nous avons tendance à considérer comme des lois générales, voire universelles, alors qu'elles ne sont que la condensation de ce qui a déjà eu lieu : elles ne peuvent prédire l'avenir qu'à la condition que celui-ci prolonge le passé, sans surprise ni invention.

Cela ouvre bien sûr des perspectives fascinantes, notamment celle de redécouvrir des lois universelles déjà connues par la simple analyse de données massives, ou bien, peut-être, d'en identifier de nouvelles. Le big data peut également aider à la compréhension de phénomènes impliquant de très grands

nombres de variables quantifiables, tels les phénomènes météorologiques ou climatiques, les comportements électoraux, l'usage des réseaux sociaux... Mais il est aussi envisageable que nous nous perdions dans l'identification de multiples corrélations, pas forcément bien interprétées, par exemple entre vie passée, vie professionnelle, activités associatives, goûts musicaux, positions politiques, relations amicales... Or, une corrélation n'est pas la même chose qu'une relation de cause à effet : ce n'est pas parce qu'il y a des grenouilles après la pluie qu'on a le droit de dire qu'il a plu des grenouilles. Mais il arrive très souvent que nous confondions les deux choses, à la manière d'un Coluche conseillant de ne jamais aller à l'hôpital au motif qu'on y meurt plus souvent que chez soi...

Le philosophe Émile Meyerson avait su exprimer le vertige qui nous saisit dès que nous voulons comprendre le rôle de la causalité dans la nature : « Les phénomènes naturels se suivent et semblent se commander, écrivait-il en 1934 : la cause entraîne l'effet. Mais comment cela se fait-il ? Quel est le lien qui relie l'effet à la cause, et y a-t-il même un lien, un ciment qui les rattache l'un à l'autre ou s'agit-il, comme Hume a cru pouvoir le démontrer, d'une pure illusion de l'esprit, née de ce que nous avons l'habitude de les voir se poursuivre dans le temps et que nous transposons cette constatation en nécessité, en transformant la *séquence* en *conséquence*¹. » Il y aurait en somme, en notre esprit, la tentation, justifiée ou non, de considérer toute séquence d'événements comme le déroulement temporel d'une suite concaténée dont le dernier événement serait la conséquence directe ou indirecte du premier. Quelque chose dans notre entendement réclamerait donc de voir le monde comme une vaste chaîne causale, comme le disait déjà un vieux proverbe latin : *Post hoc, ergo propter hoc* (sous prétexte qu'un événement précède un autre événement, on croit qu'il en est la cause).

Il y a donc bien la possibilité qu'avec le big data, au lieu de théoriser, nous cédions aux facilités de l'induction et ►►

¹ Émile Meyerson, « Philosophie de la nature et philosophie de l'intellect », 1934, *Revue de Métaphysique* 41, p. 115.

délaissions le « geste théorique », celui qui consiste à énoncer des hypothèses portant bien au-delà des données disponibles, à l'instar de Mendel postulant l'existence d'un matériel génétique bien avant la découverte de l'ADN.

Considérons un autre cas exemplaire, celui d'Einstein. En 1915, il publiait la théorie de la relativité générale, alors qu'on n'avait que très peu de données sur l'univers: on ignorait, par exemple, qu'existaient d'autres galaxies que la nôtre, on ne savait pas d'où vient que les étoiles brillent, ni que l'univers est en expansion, etc. Mais les équations d'Einstein, d'une part se sont parfaitement accommodées de la quantité gigantesque de données recueillies depuis un siècle par les télescopes et les satellites, d'autre part ont permis de prédire l'existence de nouvelles sortes d'objets physiques, tels

les trous noirs ou les ondes gravitationnelles. Imaginons maintenant que les choses se soient passées dans l'ordre inverse, c'est-à-dire que nous ayons commencé avec toutes les données dont nous disposons aujourd'hui, mais sans avoir à notre disposition la théorie de la relativité générale. Pourrions-nous, par une sorte d'induction théorique permettant de passer des données aux lois, découvrir les équations d'Einstein? La réponse de ce dernier aurait été négative, du moins si l'on en croit la lettre qu'il écrivit un jour à son grand ami Maurice Solovine: « Aucune méthode inductive ne peut conduire aux concepts fondamentaux de la physique. L'incapacité à le comprendre est la plus grave erreur philosophique de nombreux penseurs du XIX^e siècle. »

En la matière, le XXI^e siècle pourrait-il changer la donne? ■

Depuis 1976
La revue

L'INFORMATION IMMOBILIÈRE

économie
art
architecture
immobilier
environnement

DERNIER NUMÉRO



ABONNEMENTS:
Route de Chêne 36 – CP 6255
1211 Genève 6 | Tél. 058 810 33 31/24
abonnement@informationimmobiliere.ch

LES ATOUTS DE LA REVUE

- ▶ **TIRAGE CERTIFIÉ**
Imprimée à 115 000 exemplaires par numéro
- ▶ **LECTORAT**
38% de cadres moyens et supérieurs
74% d'hommes dont 90% ont plus de 35 ans
35% des lecteurs consacrent entre 30 et 60 minutes à la lecture de leur revue
- ▶ **DIFFUSION**
Partout en Suisse (87% sur l'Arc lémanique)
Abonnement gratuit (2015/2016: REMP 113 788)
- ▶ **CHRONIQUEURS**
Des personnalités de renom contribuent à sa renommée
- ▶ **ÉDITORIAL**
Revue de belle facture, *L'INFORMATION IMMOBILIÈRE* diffuse une information suivie, de qualité et objective
- ▶ **RÉDACTION - PUBLICITÉ**
Rédactrice en chef: Christine Esseiva – redaction@informationimmobiliere.ch
Responsable commercial: Edouard Carrascosa – edouard.carrascosa@spg.ch

Pour vous abonner gratuitement à la revue *L'INFORMATION IMMOBILIÈRE*, nous vous remercions de nous contacter:
Route de Chêne 36 – CP 6255 – 1211 Genève 6 | Tél. 058 810 33 31/24 | abonnement@informationimmobiliere.ch
Vous pouvez aussi consulter notre site internet: www.informationimmobiliere.ch