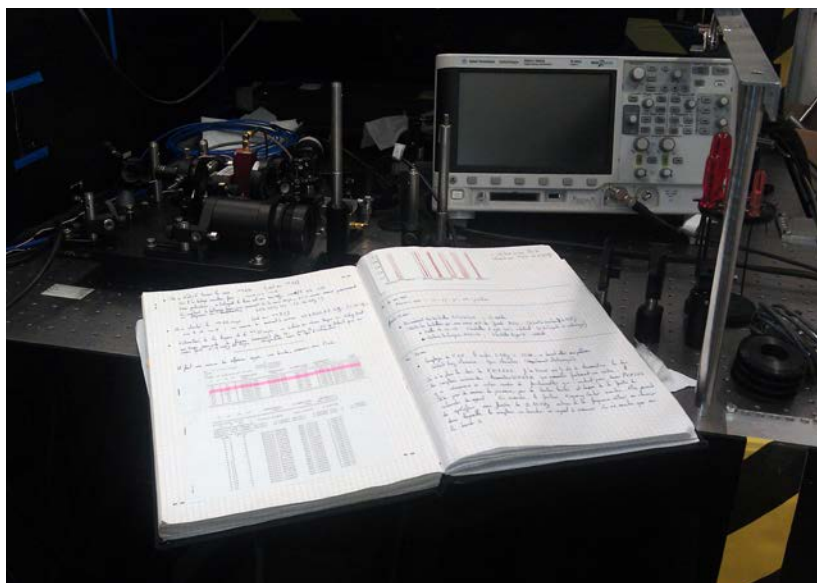




Un cahier de laboratoire, la mémoire d'une expérience, où l'on consigne les résultats. De plus en plus, ces cahiers prennent une forme électronique. Certains d'entre eux contiennent des téraoctets de données, voire plus ; leur exploitation ne peut se faire qu'avec des moyens informatiques.

**IL EST DEVENU
IMPOSSIBLE TANT
DE CONFRONTER
« À LA MAIN »
LES THÉORIES
AUX OBSERVATIONS
QUE DE S'INSPIRER
DE CES OBSERVATIONS
POUR IMAGINER
DES THÉORIES**

Cette transformation progressive de la démarche scientifique en induit une autre sur la connaissance scientifique elle-même. Jusqu'à la fin du xx^e siècle, les théories scientifiques étaient exprimées par des propositions de petite taille. Par exemple, la deuxième loi de Newton (selon laquelle la force s'exerçant sur un objet est égale au produit de sa masse par son accélération) s'exprime par une proposition d'une dizaine de caractères : $\vec{F} = m d^2\vec{x} / dt^2$. Mais rien ne garantit que les théories automatiquement inférées à partir de gigantesques volumes d'observations soient petites. De fait, nous commençons à prendre conscience qu'il n'y a aucune raison *a priori* pour que les lois de la nature ou les règles de la grammaire d'une langue s'expriment par de courts énoncés.

La grammaire d'une langue, par exemple, contient sans doute des milliers de règles dont certaines ne portent que sur des expressions très particulières. En français, on peut dire « je viendrai lundi 2 septembre » ou « je viendrai le lundi 2 septembre » ; on peut aussi dire « je viendrai demain », mais pas « je viendrai le demain ». Il semble ainsi y avoir des

règles de grammaire locales propres aux dates, qui diffèrent des règles de grammaire relatives aux adverbes de temps. Et si nous mettons bout à bout toutes ces règles de grammaire locales propres aux dates, aux noms de pays et autres toponymes (on va « en Russie », mais « au Brésil » et « à Madagascar »), on obtiendra probablement des connaissances qui ne peuvent pas s'exprimer, comme la deuxième loi de Newton ou les équations de Maxwell régissant les champs électromagnétiques, en quelques dizaines de caractères (voir l'encadré page suivante).

QUE DEVIENT LE SCIENTIFIQUE ?

En fait, cette idée est légèrement antérieure au déluge de données. Elle est sans doute apparue quand nous avons pris conscience que les lois qui régissent l'évolution de l'atmosphère ou le fonctionnement d'une cellule, par exemple, ne pouvaient pas s'exprimer en quelques lignes.

Lorsqu'une théorie s'exprime par des énoncés ou des algorithmes qui font des centaines de pages, le lien entre prévision et explication se distend. Ainsi, nous avons pris l'habitude d'accepter que les prévisions météorologiques indiquent le temps du lendemain avec une probabilité de succès assez élevée, mais qu'elles ne nous expliquent pas pour quelles raisons ce lendemain sera ensoleillé ou pluvieux. Les lois de Newton, au contraire, nous permettent non seulement de prédire que les trajectoires des planètes sont des ellipses, mais aussi de comprendre pourquoi.

Si nous ne pouvons pas expliquer le temps du lendemain, c'est précisément parce que la situation météorologique n'est pas le résultat d'une cause unique, comme le mouvement d'une planète est le résultat de l'attraction d'une étoile, mais de milliers de causes particulières desquelles il est difficile de faire émerger une explication globale.

Comme les données, ces théories sont trop complexes pour être utilisées « à la main » : seul un ordinateur peut les exploiter pour faire des prédictions. En extrapolant – à nouveau, peut-être dangereusement – les évolutions de la forme des théories scientifiques, nous pouvons imaginer que l'élaboration des prédictions sera un jour entièrement déléguée aux ordinateurs, et que les scientifiques auront alors pour rôle de concevoir des algorithmes permettant aux ordinateurs de calculer des prédictions à partir d'une théorie.

Il est cependant essentiel de garder en tête que, comme les autres théories scientifiques, ces théories restent des hypothèses, toujours à la merci d'une réfutation par les observations, et que leur pertinence ne tient qu'à leur capacité à résister à ces réfutations potentielles. Le risque est grand de prendre pour des oracles

➤ des prédictions produites par des ordinateurs à partir de théories trop complexes pour être comprises. Par exemple, contrairement à ce que l'on entend parfois, ce n'est pas le fait que tous les modèles climatiques s'accordent dans leurs prédictions qui importe, mais le fait que les prédictions passées aient été en accord, ou non, avec les observations.

Comme souvent, ces transformations de la technique et de la science en annoncent d'autres, dans divers autres domaines de notre culture. Un premier changement, conséquence de cette capacité nouvelle de collecter, stocker, transformer et transmettre de grandes quantités d'information, est celui de notre rapport à la mémoire et au passé.

La plupart des quelque 100 milliards de femmes et d'hommes qui ont vécu avant nous n'ont laissé aucune trace de leur passage sur Terre. Par exemple, malgré l'écriture et l'imprimerie, peu d'entre nous peuvent citer le nom et le prénom de leurs 16 arrière-arrière-grands-parents. Et les plus anciennes personnes dont nous connaissons le nom, sans doute quelques souverains de Mésopotamie ou d'Égypte, remontent à un passé relativement récent. De l'histoire de l'humanité durant les centaines de milliers d'années qui les ont précédés, depuis l'apparition des premiers *Homo sapiens*, nous n'avons presque aucune trace.

UN BIG BROTHER HYPERMNÉSIQUE

Les femmes et les hommes des siècles prochains auront une vision bien différente de leur passé, car les capacités actuelles de stockage de l'information rendent l'humanité hypermnésique. Ainsi, nos descendants connaîtront mieux que nous leur arbre généalogique; et au lieu d'un nom laissé, pour toute trace, dans un registre d'état civil, ils pourront conserver sur chacun de leurs lointains ancêtres une grande quantité d'informations. Et les futurs historiens rendent jaloux ceux d'aujourd'hui, car ils disposeront de documents bien plus nombreux et complets...

Cette médaille a son revers: il n'est pas si facile de vivre en sachant que chacun de ses gestes laisse potentiellement une trace à la postérité – c'est vivre avec un *Big Brother* hypermnésique. Aussi est-il probable, et souhaitable, que se développera un véritable droit à l'oubli, afin de permettre à chacun de décider des traces qu'il souhaite effacer ou transmettre. Ces préoccupations sont à l'origine des recherches des informaticiens qui développent des techniques pour protéger la vie privée, par exemple celle des personnes enregistrées dans des bases de données.

D'autres éléments de la culture humaine transformés par cette capacité nouvelle à

collecter, stocker, transmettre et transformer de grandes quantités d'information sont les institutions politiques. Notre démocratie représentative a été inventée aux XVII^e et XVIII^e siècles, à une époque où le débit de l'information que chaque citoyen pouvait communiquer à la collectivité – et, de façon symétrique, la collectivité à chaque citoyen – était très faible.

De fait, en participant tous les cinq ans à l'élection d'un chef d'État parmi 16 candidats, chaque citoyen communique à la collectivité une information de quatre bits. Le débit d'information est donc de 0,8 bit par an – soit 2×10^{-8} bit par seconde. On peut ajouter l'information communiquée lors des élections législatives, locales..., l'ordre de grandeur reste le même: quelques bits par an.

Aussi, jusqu'au XX^e siècle, de nombreuses stratégies – tracts, pétitions, manifestations, etc. – ont été inventées pour accroître cette quantité. Ces formes de militantisme ont vieilli et sont aujourd'hui remplacées par d'autres: sites Web, pétitions en ligne... La faible quantité d'information échangée a également pour conséquence que, lors d'une unique élection, nous décidons à la fois du devenir de l'armée, de la police, de la justice, de la diplomatie, de l'éducation, de la recherche, de la santé, des arts, du logement, de

DES THÉORIES SCIENTIFIQUES QUI S'ALLONGENT ?

Les phénomènes électriques et magnétiques, telles des étincelles (à gauche), sont bien décrits à partir de la théorie électromagnétique élaborée par James Clerk Maxwell vers 1870. Cette théorie s'exprime sous forme compacte par quatre équations vectorielles (les « équations de Maxwell ») mettant en jeu le champ

électrique $\vec{E}$ et le champ magnétique $\vec{B}$. En revanche, les prévisions météorologiques (à droite) résultent de modélisations théoriques et de calculs qu'il est impossible d'exprimer sous forme courte. Avec les données massives, les situations de ce type seront probablement de plus en plus fréquentes.

