

Hubert Guillaud

Un monde de données

comprendre l'implication sociale et politique
des banques de données et leur accès

PUBLIE PAPIER

Table

Présentation	15
--------------	----

Première partie

Vers un monde de données en réseaux	21
Est-ce que le déluge de données va rendre la méthode scientifique obsolète?	21
Demain l'intelligence des données	26
L'histoire de l'innovation contemporaine c'est les Big Data	31
Quand nos requêtes aident à tracer les évolutions de la grippe	36
Nous sommes le Web!	38
Vers le web implicite	41
Le siècle des réseaux	45

Deuxième partie

Le Web ²	69
Le Web à la puissance 2: le Web 2.0 cinq ans plus tard	69
Redéfinir l'intelligence collective: une nouvelle participation sensorielle	74
Faire coopérer les sous-systèmes de données	78
Comment le Web apprend : sens explicite contre sens implicite	81

Quand le Web rencontre le monde :	
« l'ombre informationnelle » et « l'internet des objets »	84
L'essor du temps réel: un cerveau collectif	94
En conclusion: ce qui compte	99
Web puissance 2: quelques exemples	101

Troisième partie

Quelles données?	111
Quelles données libère-t-on?	111
Il n'existe pas de données brutes	113
Taxonomie des données sociales	115
Mesurer les résultats de la libération des données	118
Quelle est la valeur de la libération des données?	118
Quels critères président à l'ouverture des données	120
L'abus de données n'est pas bon pour la vie privée	121
Toutes les données sont devenues personnelles	122
Peut-on réellement anonymiser des données?	125
Quelles intimités le partage libère-t-il?	128
De quelle transparence ces données sont-elles porteuses?	129
La transparence a-t-elle des limites?	132
Peut-on être contre la transparence?	133
Tyrannie de la transparence ou tyrannie de la vertu?	135
Les données ne prouvent rien, mais l'argent semble tout prouver	138
Comment éviter la démocratie du cynisme?	140
La transparence est-elle de droite ou de gauche?	142

Ce que menace Wikileaks, c'est le fonctionnement même du pouvoir ; ce qui disparaît, c'est l'espace public	145
Wikileaks: l'apologie du modèle conspirationniste?	146
Le paradoxe de l'espace public	148
La désintégration des apparences	151
À défaut de réduire la collecte des données, comment les altérer?	152
Les limites de la fouille sociale des données	155
Que faire face à la puissance des données?	159
Les illusions de la maîtrise de l'identité	160
Les politiques de confidentialité doivent être au cœur de nos systèmes socio-techniques	162
Si les données sont le code, si le code c'est la loi, alors je veux avoir accès aux données	167
Ouvrir, ou tout ouvrir	168
Vie privée: où sont les régulateurs? Où sont les régulations?	171
Pourquoi n'avons-nous pas le choix?	171
Où est la régulations? Où sont les régulateurs?	175
De nouveaux droits à conquérir	178

Quatrième partie

Des données aux Big Datas	185
L'avenir de l'information: le Web de données	185
Big Data: les progrès de l'analyse des données	190
Vers le commerce algorithmique	191
Les raisons de la révolution	192

Sciences sociales, les données ne feront pas tout	198
Big Data: la nécessité d'un débat	203
L'automatisation de la recherche change la définition du savoir	211
Les revendications d'objectivité et d'exactitude sont trompeuses	216
De plus grosses données ne sont pas toujours de meilleures données	220
Toutes les données ne sont pas équivalentes	227
Accessible ne veut pas dire éthique	232
L'accès limité aux Big Data crée de nouvelles fractures numériques	237

Cinquième partie

Enjeux	247
Du discours aux données: vers la fin de la rhétorique?	247
L'open data: la donnée comme alternative au discours	248
Quand la donnée produit son propre discours	251
Storytelling: la sacralisation du discours au détriment des faits	252
La mise à disposition des données réinterroge leur nature	255
La capacité prédictive de nos systèmes socio-techniques va-t-elle tuer notre livre arbitre?	257
Quelles informations retenir? Qui va les retenir pour nous?	258
Les algorithmes prédictifs vont-ils décider pour nous?	260

Kevin Slavin: « il nous faut dresser l'Atlas des algorithmes contemporains »	266
Nos vies gérées par les données	273
Je me mesure, donc je suis	275
La montée des capteurs de soi	277
Maîtriser les machines qui nous mesurent	280
L'insupportable objectivité des machines	283

Bibliographie / Remerciements

*À nous, internautes, qui seront les maîtres (ou pas)
de ce monde de données...*

Présentation

NOUS SOMMES ENTRÉS dans un monde de données. Un monde où ces données sont en passe de devenir l'essence même de la connaissance et de l'information. La donnée (*data*) est la plus petite part d'information accessible, à la manière des zéros et des uns qui constituent les bits d'information à l'heure du numérique. Elle est déterminée par des valeurs, par des champs qui s'appliquent à elle et la quantifient.

Avec l'informatisation de notre quotidien, tout devient données. Elles sont le moteur du Web sémantique* de Tim Berners-Lee (ce Web de données* ou Web 3.0*) comme du Web²* de John Battelle et Tim O'Reilly. C'est le moteur du Web implicite*, celui qui comprend le moindre de nos clics. C'est bien sûr le moteur principal du Web 2.0* et des interfaces de programmation qui le structurent*. C'est aussi

* lien hypertexte présent dans la version numérique, disponible à l'adresse suivante : <http://www.publie.net/fr/ebook/9782814505063/un-monde-de-donnees> (version numérique gratuite incluse, voir en dernière page). Les liens HTML seront présentés ainsi tout au long de cet ouvrage.

celui du Web relationnel*, qui fait communiquer les données entre elles. C'est le moteur de la compréhension de nos existences* et de nouvelles perspectives scientifiques*, journalistiques*, marketing ou démocratiques*.

Ces données deviennent intelligentes* par leur recouplement et leur capacité à adapter notre environnement au contexte que les machines perçoivent de nous via leurs capteurs. Ce sont elles qui nous font entrer dans le Siècle des réseaux*, qui s'apprêtent à transformer la connaissance, jusqu'au fondement de la méthode scientifique*.

Ces données sont multiples, hétéroclites, hétérogènes, mais elles se relient les unes aux autres. Elles répondent à des principes*, des taxonomies*, et produisent chaque jour des effets* toujours plus puissants sur le corps économico-social de notre société.

L'un de leurs plus formidables espoirs repose sur leur capacité à modifier le réel et plus encore sur leur capacité prédictive. « Au risque de tuer notre libre arbitre, individuel comme collectif », comme l'explique Antoinette Rouvroy, chercheuse au Fonds National de la Recherche Scientifique (FNRS) dans l'émission que consacrait à ce sujet le dernier numéro de Place de la Toile.

Les menaces de cette société de la donnée sont à la hauteur de leur puissance*, de cette nouvelle compréhension

de l'individu et de la société que les données impliquent. Nous sommes entrés dans un monde où notre vie privée est désormais en réseau*, où toutes les données sont potentiellement personnelles*. Un monde où l'alternative n'est déjà plus de les contenir mais de trouver les moyens de les altérer pour préserver sa vie privée*, quand bien même leurs promesses ne seraient pas toutes tenues*...

Car le monde qu'elles contribuent à façonner n'est pas encore aussi lisse et lisible qu'elles le promettent. Mais pour combien de temps encore ?