

# Sens [public] Revue Web

## CRÉATIONS

Marcello Vitali Rosati

CLASSER CALCULER

[| Se connecter](#)

4 mai 2012



### **Le rêve d'un savoir total, de l'information à la donnée brute**

Un rêve, plus ou moins explicite, hante nos esprits depuis plusieurs millénaires. On le retrouve ci et là dans les listes égyptiennes, dans les catalogues aristotéliens, dans les règles mnémotechniques des néoplatoniciens florentins de la renaissance, dans les constructions mathématiques de Leibniz, dans les affirmations des grands noms du web : le monde est constitué d'une masse énorme d'informations, dont la connaissance et l'exploitation permettrait la maîtrise quasi-totale. Il serait alors possible de tout savoir, de tout prévoir, de tout faire. Mais deux limites, proprement humaines, empêchent la détention et l'exploitation de cette globalité d'informations : l'accessibilité et la calculabilité.

Les informations ne sont pas toutes accessibles à l'homme, même en considérant qu'elles existent, qu'elles résident quelque part – contenues dans un livre, détenues par un ensemble de spécialistes ou qu'elles soient simplement observables dans la nature – elles sont, ou bien disséminées, ou bien cachées, ou bien incompréhensibles, et ce en raison de l'incapacité ou nous nous trouvons à déchiffrer et uniformiser les langues, des codages, des formats. Le problème structurel de la bibliothèque dérive de ce constat. La bibliothèque est un dispositif qui essaie d'abattre l'une des deux limites du rêve de savoir total : l'accessibilité. Dans une bibliothèque toutes les informations – ou du moins beaucoup d'informations – sont accessibles. Mais, quand bien même l'accessibilité serait totale, il resterait une autre limite : la concaténation des données, la calculabilité. En d'autres termes, une fois qu'un homme a accès à une grande masse d'informations, comment peut-il les exploiter ? Ayant accès à des milliers de livres, je ne peux exploiter que les informations que je suis capable de lire et de retenir. Pour pouvoir exploiter ces informations il faudrait qu'elles constituent un objet calculable par une immense machine qui les exploiterait automatiquement. La mathématisation des sciences est une tentative de résolution de ce second problème.

La ferveur suscitée par le développement du web naît d'un tel rêve et promet une accessibilité ainsi qu'une calculabilité totales. La machine devient le dieu leibnizien, capable de tout savoir et de tout calculer, accessibilité totale et calculabilité totale signifiant savoir total. Tim Berners Lee peut être considéré comme un des promoteurs du rêve de savoir total[1]. Dans un TED Talk de 2009[2], il présente cette idée avec un ton enthousiaste et enthousiasmant. « Il y a quelques années, je vous ai demandé de mettre vos documents en ligne et vous l'avez fait. C'est génial ! Aujourd'hui je vous demande de mettre vos données en ligne. » Accessibilité et calculabilité. Le web a permis de rendre accessible une masse énorme de documents. Maintenant le problème est : comment les exploiter ? Un homme ne peut pas les utiliser à cause de leur masse. Mais si les informations deviennent des données pures, alors les machines peuvent les comprendre et les calculer : le savoir total est donc possible. « Raw Data Now ! » s'écrie Tim Berners Lee : des données brutes tout de suite. La donnée brute est une formalisation de l'information qui rend la connaissance exploitable par une machine. L'analyse des données peut être réalisée par la force de calcul de l'ordinateur. La donnée est pure, ouverte à toutes les associations de sens possible et une machine peut la relier à l'ensemble des autres données. Le monde peut être connu de façon totale, absolue et objective.

Si l'on admet que le web d'aujourd'hui a résolu le problème de l'accessibilité à l'information, que tout est en ligne et potentiellement accessible, le web de demain, le web des données ou web sémantique, résoudra le problème de la calculabilité.

### **Les errements du sens, le choix de la liste**

On ne peut nier que cette idée soit attirante, mais une nouvelle limite se présente : la calculabilité ne résout pas la question de la navigation, du parcours, du sens.

En admettant que toutes les informations soient potentiellement accessibles, que le monde soit sur le web, il en va du web comme de la bibliothèque : si le monde est devenu une bibliothèque, quel bibliothécaire sera capable d'en parcourir les rayons et de trouver le livre, la construction de sens précisément recherchée au moyen d'une requête. C'est ici que nous retrouvons le paradoxe de la Bibliothèque de Babel imaginée par Borges[3], bibliothèque contenant tous les livres possibles, y compris ceux qui ne présentent que des combinaisons de signes dépourvues de signification. La quantité de livres présents dans la bibliothèque est finie mais très élevée. Le bibliothécaire de Borges ne dispose pas d'une grande puissance de calcul. Or, même s'il l'avait, cela ne lui servirait à rien, car le problème n'est pas de pouvoir recenser toutes les données, mais de leur donner un sens. En matière d'information, la donnée pure ne signifie tout simplement rien. Elle n'a de signification qu'en relation avec un parcours, un chemin de lecture. Chaque parcours ou suite de liens est aléatoire et le nombre de parcours possibles est considérable. Or, c'est précisément le fait d'en choisir un qui produit du sens. Le sens recherché a tendance à se perdre dans le nombre vertigineux des parcours possibles. Il nous faut alors avoir recours à un principe de raison pour démêler tous les nœuds du possible, pour choisir un possible parmi les autres. La puissance de calcul, représentant la capacité de l'ordinateur à prendre en compte un maximum de données en un minimum de temps, peut produire pour nous des listes de données répondant à une demande. Au sein de chaque liste, les données changent de signification. Or comment choisir la liste correspondant à notre recherche ?

Nous avons tendance à nous en remettre au calcul automatique : « I've got a problem but I feel lucky. » Tel est le concept du bibliothécaire du web le plus consulté : Google. Sur la masse exorbitante des parcours possibles, il y en a un qui est le bon et si j'ai de la chance, Google va me le donner. La bonne liste est là, avec son agencement de données unique et son sens unique. M'entraîne-t-elle idéalement sur cette voie que je cherchais sans la connaître, me conduit-elle sur le chemin que je m'attendais à emprunter dès le début, ou m'oriente-t-elle vers une toute autre destination pour m'imposer une autre construction de sens ? A-t-elle seulement un sens ?

De surcroît, une liste, une série de relation entre données ne revêt de signification qu'à l'instant et dans le contexte où nous la parcourons, suivant la grille de lecture que nous adoptons dans le contexte où ce parcours s'effectue. La machine ne produit pas de sens. Le calcul des possibles n'est pas un dispositif signifiant. Dans ce cas, le dieu leibnizien est parfaitement incapable de choisir le meilleur des mondes : le principe de raison se transforme en une question de sens et le sens affiche de plus en plus sa partialité. Cela implique de considérer la particularité des parcours qui produisent un sens bien différent selon les contextes de chaque traversée, le possible calculable étant ainsi davantage à considérer comme un virtuel à parcourir.

[1] On peut considérer le web comme la tentative de réaliser ce rêve de savoir total. Cf. l'article écrit en 1945 par Vannevar Bush (*As we may think*) qui a inspiré le développement de l'hypertexte. [Disponible en ligne](#).

[2] [Voir en ligne](#).

[3] Borges, *Fictions*, Paris, Gallimard, 2009 p. 72.

[Répondre à cet article](#)