

Big Data, Big Bêta... ou Bide Data ?

Par Sami Coll, Sociologue à l'Université de Lausanne, 31 janvier 2016, samicoll.blog

Comme Pierre Richard à quatre pattes devant le ragoût de son toutou dans « Le coup du parapluie », le numérique, j'en suis fou. Vraiment. Depuis que j'ai effleuré le clavier de l'un des premiers ordinateurs grand public en 1980. Si bien que comme Gainsbarre, « Quand on m'dit que j'suis [technophobe], j'me marre doucement pour n'pas te réveiller », ma p'tite numérique. Il est vrai que depuis un certain temps, un moulin à prière s'est mis à tourner de plus en plus vite : numérique, numérique, numérique ! Il faut l'aimer, le chérir, le suivre, lui obéir. Et s'il on n'obtempère pas, il nous écrasera. A l'impératif « tu te soumettras à l'économie », réputée objective et essentialiste, succède « tu te soumettras au big data ». Si bien que l'on oublie que le deuxième se base sur le premier. Le big data, dernier mot valise pour caractériser la révolution numérique, nous aura fait oublier sur quoi il repose : un ubercapitalisme féroce, longuement mijoté, qui remet en cause le progrès social et intellectuel durement acquis depuis des siècles.

A ce big machin, on lui prête à terme la plus grande des vertus : l'intelligence absolue. Quand on aura enfin tout récolté, analysé, comparé, mesuré, on pourra produire une connaissance universelle qui nous libérera des maux de l'humanité. Pourtant, même les « data scientists » les plus brillants sont les premiers à nous expliquer que les données sont résistantes. Elles peinent à livrer le jus de la vérité. Ces tenants d'une nouvelle profession en vogue font la distinction entre « donnée brute », « information » et « connaissance », et reconnaissent volontiers que le passage d'une catégorie à l'autre est des plus complexes. Mais si l'idée de « donnée brute » était déjà une aberration ? Ontologiquement, une donnée est forcément le produit d'un contexte géographique, social, historique et politique. Pourquoi, pour prendre l'exemple que j'ai longuement développé dans mon livre « Surveiller et récompenser », la Migros s'intéresse-t-elle à la donnée produite par l'interaction entre un terminal technique, le lecteur de carte Cumulus de la caisse, et la carte d'un client, plutôt que par celle qui quantifierait la bave produite par les escargots de bourgogne ? La première est le résultat d'une attente préparée. Il aura fallu installer un terminal, le connecter, produire des millions de cartes en plastique dotée de codes-barres distincts, faire en sorte que les clients les portent sur eux et adoptent le réflexe de la présenter à chaque injonction. La deuxième servira peut-être à enrichir une connaissance fondamentale portée par des biologistes talentueux. Je ris encore de cette publicité télévisuelle nocturne pour de la crème miracle à base de bave d'escargot. Il n'y a pas grand-chose de neutre dans ces chiffres. Certes, il devient parfois difficile de retracer l'histoire des intentions qui ont motivé le développement d'un dispositif de captation et de traitement des données. Autre exemple. A ses débuts, Google a mis au point un algorithme et des serveurs dans l'objectif de proposer un moteur de recherche simplifié mais redoutablement efficace pour offrir une alternative à la saturation des publicités sur les moteurs vedettes d'alors : Altavista, Yahoo, Excite... Aujourd'hui, il enrichit sa puissance de calcul pour s'enrichir tout court, afin aussi, si on veut être optimiste, de rendre service à la société civile. « Don't be evil », comme ils disent. Dans tous les cas, que les intentions soient bonnes ou mauvaises, elles ne sont en tout cas aucunement neutres.

Et c'est sur ce point qu'il est urgent de se réveiller. Le big data n'est pas un capital de données flottantes – on entend parler aussi de « nouvel or noir » – préexistant aux objectifs de leur extraction pour fabriquer un monde digital censé représenter la réalité. Il est fondamentalement biaisé par une conception du monde, du social, des rapports sociaux et des attentes politico-économiques qui financent également son développement. Et celles-ci sont franchement influencées par un capitaliste informationnel davantage au service de l'efficacité et du profit que celui de la recherche fondamentale ou de valeurs telles que l'égalité des chances. C'est une version du projet du digital

auquel je n'adhère personnellement pas. Je respecte toute opinion contraire argumentée et assumée, mais je trouve qu'il est devenu malhonnête et inacceptable de partir du principe que le big data reposerait sur quelque chose de « naturel », vrai en soi et, surtout, inéluctable. Ses interstices politico-techniques doivent être soigneusement étudiés et discutés, afin de pouvoir reprendre la barre du navire politique. Les enjeux éthiques d'un tel bouleversement doivent s'inviter de toute urgence dans les discussions politiques, aujourd'hui bloquées par une admiration béate des miracles du numérique qui n'alimente qu'une recherche des conditions de réalisation de cet « écosystème socio-technique ».

Dans un sens, le big data est encore en friche. Big bêta, pourrait-on même l'appeler. Bêta comme la lettre grecque et le suffixe que l'on trouve après le nom des logiciels en première phase de test. Mais bêta, aussi, comme pourrait le devenir notre société qui ne se pose pas les bonnes questions et se laisse laminer par les tenants d'un nouveau pouvoir politique. Je parle de celui qui est d'ores-et-déjà exercé par les géants du numérique avec leurs boîtes noires algorithmiques encore plus opaques que la Stasi à son apogée. Le terme pig data me traverse parfois l'esprit. Honoré par l'invitation de Myret Zaki à contribuer à la vie des blogs de bilan.ch, je contribuerai de manière régulière pour ajouter mes propres bêtises à celles de mes estimés confrères de tous bords que je souffre fréquemment de voir aveuglés par la lumière blafarde du tout numérique. J'espère ainsi contribuer au débat de manière constructive, pour éviter, peut-être, qu'on se prenne un bide data majestueux, tant social qu'économique.