

## « Data, nouvelles approches de traitement électronique de l'information et bien commun »

Mots clefs : DATA, DONNEES, WEB, RESEAUX SOCIAUX, BIEN PUBLIC, MEDIAS, BIEN COMMUN, NOUVELLES TECHNOLOGIES, INTERNET.

Les capacités de calcul quasiment illimitées développées par les ordinateurs permettent, en compilant les données en tout genre récoltées sur internet ou via différents capteurs, de produire une information radicalement nouvelle. Si les médias peuvent à coup sûr espérer bénéficier de ces nouvelles sources pour réinventer leur activité au service du bien commun, ils peuvent aussi redouter d'être dépassés par des technologies qu'ils ne maîtrisent pas et par des acteurs qui ne remplissent pas les mêmes missions qu'eux.

### Henri Pigeat

Les diverses formes de capteurs qui enregistrent le moindre de nos mouvements produisent une masse de données (*data*) sans équivalent dans l'histoire, qui ne peut être traitée que par ordinateur et produit une information d'un type nouveau. La conséquence en est une redéfinition des rapports de forces dans un paysage médiatique en cours de renouvellement profond et peut être l'amorce d'une mutation quasi anthropologique. L'information du public en ressort plus riche, mais l'accès à la vérité se trouve aussi en quelque sorte déréalisé dans un univers où c'est moins le fait qui importe que la collection de faits. De surcroît, le croisement de données de nature diverse pour profiler les publics tend à segmenter la société tout en menaçant l'existence même d'un espace privé. La dépendance accrue des journalistes comme de l'ensemble des acteurs de ce nouveau système aux ordinateurs et à leurs méthodes de calcul impose de repenser leur formation et de s'interroger en profondeur sur la façon dont on interroge la machine. L'illusion de la neutralité de la machine, de même que son incapacité à émettre des jugements moraux ou à faire preuve d'intuition rendent plus que jamais nécessaire son utilisation raisonnée par des professionnels de l'information.

### Alexandre Monnin

La possibilité de collecter des données librement accessibles sur Internet pour les compiler et en tirer des informations valorisables d'un point de vue commercial ouvre des difficultés analogues à celles posées par la tragédie des communs. Le cas de Wikipedia en offre une illustration : l'encyclopédie coopérative produit une information riche et référencée quoique non originale, qui peut être compilée dans le cadre du web des données voire valorisée à des fins commerciales. C'est par exemple ce que fait Google qui propose en réponse à de nombreuses recherches une division de la page de réponse entre une colonne de réponses classique, où l'ordre des réponses peut être modifié par les plus offrants, et le *Knowledge graph*, qui apparaît à droite et reprend des données récupérées et réordonnées à partir de Wikipedia.

Plus généralement, la valorisation commerciale de données disponibles gratuitement sur Internet pose des problèmes économiques (comment appréhender les externalités positives de la mise à disposition des données ?) juridiques (comment contrôler voire taxer l'usage de ces données ?) et plus généralement un problème de société sur la place qui doit être celle des citoyens qui font l'objet des données.

## Discussion

- Le ciblage de plus en plus fin des messages publicitaires et des liens recommandés sur les sites d'information permet d'alléger l'impact négatif de la publicité en l'adaptant à celui qui la reçoit ; mais il fait perdre une des richesses de l'action de feuilleter un journal, qui faisait rencontrer des informations qu'on n'aurait pas crue faite pour soi.
- La question des modes de régulation de l'utilisation des données a fait l'objet de nombreuses discussions : n'est-il pas envisageable que l'ONU qui a déjà créé des agences ad hoc pour faire face à l'émergence de nouvelles questions prenne en main ce sujet qui ne peut avoir de réponse qu'internationale ? Il importe surtout de voir qu'il existe des intermédiaires entre une régulation étatique difficile à rendre acceptable par les acteurs de l'internet, et une autorégulation du marché qui ne met aucune bride aux grands acteurs du web.
- La question de la propriété des données a été soulevée : le droit français ne protège pas les données mais la structure des bases de données, ce qui accentue le pouvoir des grands acteurs du web qui peuvent s'approprier les données collectées par d'autres dans des bases restructurées. Le fait que les citoyens, qui sont à la source des données, soient très peu pris en compte dans cet enjeu de propriété demeure problématique.