

Zeitschrift: Horizons : le magazine suisse de la recherche scientifique
Herausgeber: Fonds National Suisse de la Recherche Scientifique
Band: 31 [i.e. 30] (2018)
Heft: 119: La métamorphose de la Big science : comment les mégaprojets de recherche se sont ouverts à d'autres disciplines

Artikel: Un libre marché de la donnée
Autor: Wüstholtz, Florian
DOI: <https://doi.org/10.5169/seals-821648>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 02.02.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Un libre marché de la donnée

Les données sur Internet doivent être accessibles et interprétables par les algorithmes. Mais comment convaincre les géants du Web de partager leurs précieuses informations? Une équipe zurichoise propose un marché des données. *Par Florian Wüstholtz*

Le Web des données repose sur une vision assez simple: les informations en ligne sont systématiquement annotées par des labels structurés et standardisés afin de les rendre interprétables par les machines. Les algorithmes peuvent alors exploiter les contenus Internet de manière efficace et précises. La recherche et le secteur public mettent déjà à disposition de nombreuses données, notamment statistiques. Pour convaincre les prestataires commerciaux de rejoindre ce mouvement, des informaticiens de l'Université de Zurich proposent une ébauche d'un marché des données qui offrirait les bonnes incitations économiques.

«Les clients ne paient que pour ce qui les intéresse vraiment.»

Tobias Grubenmann

Les informations référencées selon les principes du Web des données sont compréhensibles pour les machines: un algorithme peut ainsi reconnaître si le nombre figurant dans un tableau indique l'âge d'une personne, une somme d'argent ou la hauteur d'un immeuble. Il devient alors possible de mettre en relation des données a priori isolées et de déterminer de nouvelles corrélations. Autant d'informations pertinentes et de qualité à disposition des milieux scientifiques, des autorités, des entreprises ainsi que de la société civile.

«Sur un tel marché, tous les prestataires imaginables seraient en mesure de vendre leurs jeux de données», explique le premier auteur de l'étude, Tobias Grubenmann de l'Institut d'informatique de l'Université de Zurich. Les autorités touristiques pourraient par exemple vendre des statistiques précises sur les nuitées qui, combinées avec des mesures des flux de circulation, permettraient de définir de nouvelles stratégies pour un tourisme durable.

Mais cette vision pose un problème: le cœur du modèle d'affaires de sociétés telles que Facebook ou Google se base sur l'accumulation des données générées par les internautes. Elles n'ont en principe aucun intérêt à partager gratuitement ce trésor. Car la monétisation du Web des données reste encore largement inexplorée

et désorganisée, souligne Tobias Grubenmann: «De nombreux prestataires vendent des bases de données dont les contenus sont très peu pertinents, et les usagers potentiels ne savent souvent pas ce qu'ils peuvent vraiment en faire. Le marché que nous proposons offre une piste sur la manière dont les fournisseurs pourraient mieux commercialiser leurs données.»

Les données au kilo

Le concept se base sur une tarification flexible déterminée par l'utilisation effective. Les données ne sont pas vendues en de coûteux paquets globaux, mais en fonction de la quantité - à la pièce pour ainsi dire. Après avoir présenté une demande, les clients peuvent discerner les éléments pertinents présents dans le jeu de données, leur qualité et leur prix d'achat. «Au final, ils ne paient que pour les données qui les intéressent vraiment», explique Tobias Grubenmann.

Dans ce système, le marché ne joue qu'un rôle d'intermédiaire. Les prestataires paient une taxe pour mettre leurs données à disposition alors que les données publiques restent accessibles gratuitement. Ce modèle constituerait une incitation à générer des données de qualité utilisables concrètement. Pour Tobias Grubenmann, il s'agirait d'un libre marché dans lequel l'offre, la demande et la réputation s'autorégulent.

L'avantage d'un tel marché est évident: lorsque les données sont clairement structurées, il est possible de les prospecter et de les relier à volonté. Les sources publiques pourraient être croisées avec les sources commerciales et générer de nouvelles connaissances. Cela ouvrirait des possibilités totalement nouvelles pour la recherche exploratoire, les visualisations ou encore les prévisions démographiques.

«Encourager la mise en lien des données publiques et de celles générées par le secteur commercial constitue quelque chose de fondamentalement positif», estime Matthias Stürmer, du Centre de recherche pour une informatique durable à l'Université de Berne. Ce marché pourrait s'avérer éminemment utile pour la recherche. Il nécessiterait toutefois un cadre légal conséquent: «La protection des données est le bien le plus précieux sur un tel marché. La respecter exige par exemple une anonymisation rigoureuse. Il faudrait également que la combinaison de différents jeux de données ne permette pas de déductions sur des particuliers.»

Florian Wüstholtz est journaliste libre à Berne.

T. Grubenmann et al.: FedMark: A Marketplace for Federated Data on the Web. arXiv preprint arXiv:1808.06298 (2018)

