

# Les mégadonnées dans la recherche

Nicolae Sfetcu

08.02.2020

Sfetcu, Nicolae, « Les mégadonnées dans la recherche », SetThings (8 février 2020), URL = <https://www.setthings.com/fr/les-megadonnees-dans-la-recherche/>

Email: [nicolae@sfetcu.com](mailto:nicolae@sfetcu.com)



Cet article est sous licence Creative Commons Attribution-NoDerivatives 4.0 International. Pour voir une copie de cette licence, visitez <http://creativecommons.org/licenses/by-nd/4.0/>.

Une traduction partielle de :

Sfetcu, Nicolae, « Etica Big Data în cercetare », SetThings (6 iulie 2019), DOI: 10.13140/RG.2.2.27629.33761, MultiMedia Publishing (ed.), ISBN: 978-606-033-228-2, URL = <https://www.setthings.com/ro/e-books/etica-big-data-in-cercetare/>

L'éthique des mégadonnées implique l'adhésion aux concepts de bons et mauvais comportements concernant les données, en particulier les données personnelles. L'éthique des mégadonnées se concentre sur les collecteurs et diffuseurs des données structurés ou non structurés.

L'éthique des mégadonnées est soutenue, au niveau de l'UE, par une documentation complète, qui cherche à trouver des solutions concrètes pour maximiser la valeur des mégadonnées sans sacrifier les droits humains fondamentaux. Le Contrôleur européen de la protection des données (CEPD) soutient le droit à la vie privée et le droit à la protection des données personnelles dans le respect de la dignité humaine. Selon ces documents, le conflit conceptuel entre confidentialité et mégadonnées, et entre confidentialité et innovation, doit être surmonté. Il est

essentiel d'identifier les moyens d'intégrer la dimension éthique dans la conception des innovations. (European Economic and Social Committee 2017)

Selon le nouveau règlement de l'UE 2016/679, les opérateurs de données doivent mettre en œuvre les mesures et technologies de confidentialité pour améliorer la confidentialité lors de la détermination des modalités de traitement et du traitement lui-même. Grâce à l'Agence de l'Union européenne chargée de la sécurité des réseaux et de l'information (ENISA), de nombreuses stratégies de confidentialité ont été identifiées par leur conception (minimisation des données, masquage des données personnelles et leurs interconnexions, traitement séparé des données personnelles, choix du plus haut niveau d'agrégation, transparence, surveillance, politique de confidentialité, problèmes juridiques).

Un moyen de base pour une coexistence pacifique entre l'exploitation des mégadonnées et la protection des données est le contrôle par l'utilisateur des données personnelles, ce qui conduit à la transparence et à la confiance entre les utilisateurs et les fournisseurs de services numériques. Comme indiqué dans l'analyse d'impact du RGPD,

« L'instauration d'un climat de confiance dans l'environnement en ligne est essentielle au développement économique. En effet, s'ils n'ont pas totalement confiance, les consommateurs hésiteront à faire des achats en ligne et à recourir à de nouveaux services, y compris aux services administratifs en ligne. Si ce manque de confiance n'est pas résolu, il continuera de ralentir l'innovation dans l'utilisation des nouvelles technologies, d'entraver la croissance économique et de priver le secteur public des avantages potentiels de la numérisation de ses services. »

Dans le cas des mégadonnées, les modèles de consentement traditionnels sont insuffisants et obsolètes. « Le consentement doit être assez détaillé pour couvrir l'ensemble des traitements et des finalités du traitement et la réutilisation des données à caractère personnel. » (European Economic and Social Committee 2017)

Un problème particulier est la *portabilité* des données, soutenue au niveau de l'UE par le CEPD dans l'avis 7/2015, (MORO 2016) où il est nécessaire de garantir le droit des citoyens

d'accéder et de corriger les données personnelles grâce à un contrôle étendu. La portabilité des données peut aider à accroître la sensibilisation et le contrôle des consommateurs en transférant des services en ligne.

CEPD estime que les données à caractère personnel devraient être traitées ainsi que d'autres ressources importantes, telles que le pétrole, où les échanges ont lieu entre les parties également informés (symétrie de l'information). En fait, le marché des informations personnelles a un caractère d'asymétrie d'information, n'étant ni transparent ni équitable, les clients ne sont pas rémunérés pour les informations personnelles qu'ils fournissent. Ainsi, la portabilité des données favoriserait un environnement plus compétitif entre les bénéficiaires de ces données, les utilisateurs pouvant choisir à qui ils donnent les données personnelles.

Une autre approche implique le *magasin* des données personnelles, avec la possibilité pour l'utilisateur d'accorder ou de retirer le consentement pour ses données personnelles. (MORO 2016) (DG Connect 2015) Le magasin des données personnelles implique un « concept, un cadre et une architecture de mise en œuvre qui fait passer l'acquisition et le contrôle des données d'un modèle de données distribué à un modèle centré sur l'utilisateur. » (European Economic and Social Committee 2017) La portabilité des données pourrait garantir cela.

Le CEPD soutient la promotion des bénéficiaires responsables et la réduction de la bureaucratie dans la protection des données, grâce à des codes de conduite, des audits, des certifications et une nouvelle génération de clauses contractuelles et de règles d'entreprise obligatoires. La responsabilité des bénéficiaires des mégadonnées implique la mise en place de politiques internes et de systèmes de contrôle conformes à la législation en vigueur, à travers des solutions intelligentes et dynamiques qui garantissent le respect des principes fondamentaux

(minimisation des données, limitation des finalités, qualité des données, traitement correct et transparent des données, conception, limitation du stockage, intégrité et confidentialité).

L'éthique des données est basée sur les principes suivants: *propriété* (les individus sont propriétaires de leurs données), *transparence des transactions* (les utilisateurs doivent avoir un accès transparent à la conception de l'algorithme), *consentement* (l'utilisateur doit être informé et consentir expressément à l'utilisation des données personnelles), *confidentialité* (la confidentialité de l'utilisateur doit être protégée), *financière* (l'utilisateur doit connaître les transactions financières résultant de l'utilisation de ses données personnelles) et *ouverture* (les ensembles de données agrégées doivent être librement disponibles).

### **L'éthique dans la recherche**

Le terme *étude des données critiques* implique que les chercheurs étudient les mégadonnées d'un point de vue critique. L'étude des données dans ce contexte implique, en plus de leur analyse, l'intégration des données dans les pratiques (connaissances), les institutions et systèmes politiques et économiques, à travers l'interaction complexe entre les données et les entités qui les produisent, les détiennent et les utilisent.

Un rapport de l'OCDE (2013) souligne que, contrairement aux normes éthiques appliquées aux données de recherche communes, dans le cas des mégadonnées : (OECD 2013)

- La collecte des données n'a pas fait l'objet d'un processus officiel d'examen éthique.
- Les règles éthiques communes ne seront pas mises en œuvre dans le cas des mégadonnées.
- L'utilisation des données de recherche peut différer de l'objectif initial.
- Les données ne sont plus conservées sous forme d'ensembles discrets.

La relation entre ceux qui fournissent les données et ceux qui les utilisent est souvent indirecte et variable. Un rapport plus récent de l'OCDE (2016) soutient que cette relation est plus faible ou inexistante, les mégadonnées limitant les capacités habituelles. (OECD 2016)

Le magasin des données est important pour l'intégrité de la recherche. Les données doivent avoir une « provenance » claire, avec des sources et un traitement connus et identifiés.

De nombreuses données qui ne sont pas spécifiquement collectées pour la recherche ont des normes différentes dans la recherche de données.

Pour certaines données, souvent avec de valeur commerciale (par exemple, les données collectées sur Twitter), leur reproduction est soumise à des restrictions légales. (UK Data Service 2017)

Les magasins de données doivent respecter des normes de transparence et de reproductibilité.

### **Bibliographie**

- DG Connect. 2015. "Study on Personal Data Stores Conducted at the Cambridge University Judge Business School." Text. Digital Single Market - European Commission. August 7, 2015. <https://ec.europa.eu/digital-single-market/en/news/study-personal-data-stores-conducted-cambridge-university-judge-business-school>.
- European Economic and Social Committee. 2017. "The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context." European Economic and Social Committee. February 22, 2017. <https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/ethics-big-data>.
- MORO, Veronica. 2016. "Meeting the Challenges of Big Data." Text. European Data Protection Supervisor - European Data Protection Supervisor. November 16, 2016. [https://edps.europa.eu/data-protection/our-work/publications/opinions/meeting-challenges-big-data\\_en](https://edps.europa.eu/data-protection/our-work/publications/opinions/meeting-challenges-big-data_en).
- OECD. 2013. "New Data for Understanding the Human Condition: International Perspectives." <http://www.oecd.org/sti/inno/new-data-for-understanding-the-human-condition.pdf>.
- . 2016. "Research Ethics and New Forms of Data for Social and Economic Research," November. <https://doi.org/10.1787/5jln7vnpxs32-en>.
- UK Data Service. 2017. "Big Data and Data Sharing: Ethical Issues." [https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing\\_ethical-issues.pdf](https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf).