

# Big Data - Aspects philosophiques

Nicolae Sfetcu

19.12.2019

Sfetcu, Nicolae, « Big Data - Aspects philosophiques », SetThings (19 décembre 2019), MultiMedia Publishing (ed.), URL = <https://www.setthings.com/fr/big-data-aspects-philosophiques/>

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/>

Le big data peut générer, par inférences, de nouvelles connaissances et perspectives. Le paradigme qui résulte de l'utilisation du big data crée de nouvelles opportunités.

L'une des principales préoccupations dans le cas du big data est que les scientifiques des données ont tendance à travailler avec des données sur des sujets qu'ils ne connaissent pas et n'ont jamais été en contact, étant éloignés du produit final de leur activité (l'application des analyses). Une étude récente (Tanner 2014) indique que cela peut être la raison d'un phénomène connu sous le nom d'aliénation numérique.

Le big data a une grande influence au niveau gouvernemental, affectant positivement la société. Ces systèmes peuvent être rendus plus efficaces en appliquant des politiques de transparence et de gouvernance ouverte, telles que les données ouvertes.

Après avoir développé des modèles prédictifs pour le comportement du public cible, le big data peut être utilisé pour générer des alertes précoces pour diverses situations. Il y a donc un retour positif entre la recherche et la pratique, avec des découvertes rapides tirées de la pratique.

A. Richterich déclare que la popularité de la surveillance de l'activité des utilisateurs a été motivée par des affirmations selon lesquelles l'utilisation (et la collecte de données avec) ces appareils amélioreraient le bien-être, la santé et l'espérance de vie des utilisateurs, et réduiraient considérablement les coûts des soins de santé. (Richterich, 2018) Pour obtenir le consentement de l'utilisateur, de nombreuses entreprises ont offert des remises aux clients qui seraient disposés à donner accès à leurs données de surveillance. (Mearian 2015) Mais il y a aussi des préoccupations concernant l'influence de ces technologies sur la société, en particulier dans les questions liées à l'équité, la discrimination, la vie privée, l'abus de données et la sécurité. (Collins 2016)

Conceptuellement, le big data doit être compris comme un terme générique désignant un ensemble de technologies émergentes. Dans leur utilisation, nous devons prendre en compte les contextes culturels, sociaux et technologiques, les réseaux, les infrastructures et les interdépendances qui peuvent avoir un sens sur le big data. Le terme « big data » fait référence non seulement aux données en tant que telles, mais également aux pratiques, infrastructures, réseaux et politiques qui influencent leurs diverses manifestations. Comprendre les mégadonnées comme un ensemble de technologies émergentes semble être conceptuellement utile, car il comprend des développements numériques activés dans la collecte, l'analyse et l'utilisation des données. (Richterich, 2018)

Dans ce contexte, Rip décrit le dilemme des développements technologiques : "Pour les technologies émergentes avec leur avenir indéterminé, il y a le défi d'articuler des valeurs et des règles appropriées qui auront du poids. Cela se produit en articulant des promesses et des visions

sur les nouvelles technologies." (Rip 2013, 192) Ainsi, les technologies émergentes sont des lieux de « normativité omniprésente » caractérisés par l'articulation de promesses et de peurs, conceptualisant une telle « normativité omniprésente » comme une approche « dans l'esprit d'une éthique pragmatique, dans laquelle les positions normatives co-évoluent. » (Rip 2013)

L'éthique pragmatique souligne que les nouvelles technologies se développent dans des sociétés où elles sont associées / dissociées de manière discursive par certaines normes et valeurs. Dans le même temps, le pragmatisme affirme que l'augmentation du grand nombre de données et de pratiques liées à la recherche n'est pas une simple question de supériorité technologique. Ils forment un champ de justification normative et de contestation.

Les néo-pragmatiques dans l'approche de l'éthique aborde les connaissances épistémologiques à travers la falsification des connaissances (scientifiques), avec des évaluations critiques des structures du pouvoir social. Keulartz et al. ont proposé une approche pragmatique de l'éthique dans une culture technologique (Keulartz et al. 2004) "comme une alternative qui combine les forces de l'éthique appliquée et des études scientifiques et technologiques, tout en évitant les lacunes de ces domaines." (Richterich, 2018) Ainsi, l'éthique appliquée est une approche efficace en termes de détection et d'expression des normes impliquées dans des actions (inter) sociotechniques ou résultant d'actions sociotechniques, mais elle n'a aucune possibilité de capter la normativité inhérente et l'agent des technologies. (Keulartz et al. 2004, 5)

Keulartz et al. considère que le manque d'évaluations technologiques normatives peut ainsi être surmonté: "l'impasse qui a émergé de ce point de vue" (c'est-à-dire les "angles morts" de l'éthique appliquée) peut être surmontée par une réévaluation du pragmatisme. (Keulartz et al. 2004, 14) Le pragmatisme éthique peut être caractérisé par trois principes communs: l'anti-fondationalisme, l'anti-dualisme et l'anti-scepticisme.

L'anti-fondationalisme fait référence au principe de falsifiabilité, étant donné que nous ne pouvons pas atteindre la certitude en termes de connaissances ou de valeurs ("vérité finale"), mais les connaissances, ainsi que les valeurs et les normes, changent avec le temps. Les valeurs morales ne sont pas statiques, mais peuvent être renégociées en fonction des évolutions technologiques.

L'anti-dualisme implique la nécessité de s'abstenir de dichotomies prédéterminées. Parmi les dualismes critiqués par Keulartz figurent l'essence / l'apparence, la théorie / la pratique, la conscience / la réalité et les faits / la valeur. L'éthique appliquée a tendance à assumer de tels dualismes *a priori*, par opposition au pragmatisme, qui souligne les interrelations et les frontières floues entre ces catégories.

L'anti-scepticisme est étroitement lié à la nécessité des perspectives situées et d'une normativité explicite, liées au fondement anti-cartésien du pragmatisme.

Dans la recherche européenne, le pragmatisme était généralement rejeté comme "superficiel et opportuniste", étant associé à des "stéréotypes négatifs", (Joas 1993) étant accusé "d'utilitarisme et de méliorisme". (Keulartz et al. 2004, 15) À la fin des années 1990 et 2000, le pragmatisme a connu un renouveau dans la recherche européenne. (Baert and Turner 2004)

L'analyse du big data d'un point de vue éthique implique deux principaux aspects interdépendants : un aspect théorique (la description philosophique des éléments soumis au contrôle éthique) et une vision pragmatique (de l'impact sur la vie des personnes et des organisations). (European Economic and Social Committee 2017)

Il y a des problèmes éthiques causés par l'intelligence artificielle, et un lien étroit entre le big data et l'intelligence artificielle et ses dérivés : apprentissage automatique, analyse sémantique, exploitation des données.

Une approche éthique passe par l'agence morale avec au moins les trois conditions de causalité, de connaissance et de choix. Selon Noorman: (Noorman 2012)

- Il existe des liens de causalité entre les personnes et le résultat des actions. La responsabilité de la personne découle du contrôle du résultat.
- Le sujet doit être informé, y compris les conséquences possibles.
- Le sujet doit donner son consentement et agir d'une certaine manière.

Le professeur Floridi, dans *La quatrième révolution*, identifie le problème moral du big data avec la découverte d'un modèle simple : une nouvelle frontière d'innovation et de concurrence. (Floridi 2014) Un autre problème associé au big data est le risque de découvrir ces modèles, changeant ainsi les prédictions.

La règle de base de l'éthique du big data est la protection de la vie privée, la liberté et la discrétion de décider de manière autonome. Il convient de noter qu'il existe une tension continue entre les besoins individuels et ceux d'une communauté.

Il est possible d'identifier plusieurs problèmes éthiques liés à l'exploitation des big data (European Economic and Social Committee 2017),

- *Confidentialité* - La limite extrême de la confidentialité est l'isolement, défini par Alan F. Westin comme "le retrait volontaire d'une personne de la société en général par des moyens physiques dans un état de solitude". Moor et Tavani ont défini un modèle de confidentialité appelé. Contrôle restreint de l'accès (RALC) qui différencie la confidentialité, la justification et la gestion de la confidentialité.
- *Réalité adaptée et bulles de filtrage* - L'application sur un serveur collecte des informations en tirant des enseignements, puis utilise ces informations pour construire un modèle de nos intérêts. Lorsqu'un système utilise ces modèles pour filtrer des informations, nous pouvons

être amenés à croire que ce que nous voyons est une vue complète d'un contexte spécifique, alors qu'en fait nous sommes limités par la "compréhension" d'un algorithme qui a construit le modèle. Les effets éthiques peuvent être multiples : certaines informations peuvent être cachées, imposer des préjugés que nous ne connaissons pas, notre vision du monde peut devenir progressivement limitée, et à long terme elle pourrait générer un certain point de vue.

- *Gestion des données après le décès* - Qu'advient-il des données d'un utilisateur décédé ? Les héritiers deviennent-ils leurs propriétaires ? Les données peuvent-elles être retirées du monde numérique ? Il y a ici des problèmes juridiques et technologiques.
- *Biais d'algorithme* - L'interprétation des données implique presque toujours certains biais. De plus, il est possible qu'une erreur dans un algorithme introduise des formes de biais. Un problème éthique est notre confiance implicite dans les algorithmes, avec des risques élevés lorsque les risques ne sont pas pris en compte en raison d'erreurs de programmation ou d'exécution des algorithmes.
- *Confidentialité vs accroître la puissance de l'analyse* - Elle se réfère à la nature émergente de l'information comme un système complexe : le résultat de données provenant de différents contextes est plus que la simple somme des parties.
- *Limitation de la finalité* - Il est très difficile, voire impossible, de limiter l'utilisation des données. La confidentialité n'est pas un bloc unique, il y a des formes subtiles de perte de la confidentialité.
- *Inertie du profil numérique des utilisateurs* - Il s'agit du sujet de la réalité personnalisée. Un modèle qui implique les intérêts d'un utilisateur est généralement basé sur le comportement passé et les informations passées. Ainsi, les algorithmes ne sont pas basés

sur l'identité réelle de la personne, mais sur une version antérieure. Cela influencera le comportement réel de l'utilisateur, poussé à maintenir ses anciens intérêts et donc à ne pas pouvoir découvrir d'autres opportunités. Si l'utilisateur n'est pas conscient de ce problème, l'influence de l'inertie sera beaucoup plus importante.

- *Radicalisation des utilisateurs, conformisme et sectarisme* - les big data peuvent se forger des opinions en utilisant des algorithmes de filtrage / recommandation, des informations, des articles et des messages personnalisés et des recommandations spécifiques d'amis. Ainsi, les utilisateurs seront de plus en plus en contact avec les personnes, les opinions et les faits qui soutiendront leur position d'origine. Cette tendance est souvent cachée aux utilisateurs de systèmes basés sur le big data, avec la tendance à développer des préjugés, allant de la conformité à la radicalisation. Il est possible de postuler la formation d'une sorte de subconscient technologique ayant un impact sur le développement de la personnalité des utilisateurs, phénomène évident dans le cas des réseaux sociaux, où la distance entre le monde réel ("physique") et Internet est fortement atténuée.
- *Impact sur les capacités personnelles et la liberté*
- *Égalité des droits entre le propriétaire des données et l'opérateur de données* - Habituellement, la personne dont les données sont utilisées n'est pas leur propriétaire légal. Par conséquent, une condition minimale est que cette personne ait accès à ses propres données, lui permettant de les télécharger et éventuellement de les supprimer.

### **Bibliographie**

- Baert, Patrick, and Bryan Turner. 2004. "New Pragmatism and Old Europe: Introduction to the Debate between Pragmatist Philosophy and European Social and Political Theory." *European Journal of Social Theory* 7 (3): 267–74. <https://doi.org/10.1177/1368431004044193>.
- Collins, Tom. 2016. "Security Fears Sparked over Wearable Technology." Mail Online. December 19, 2016. <http://www.dailymail.co.uk/~-/article-4049154/index.html>.

- 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>.
- Floridi, Luciano. 2014. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. OUP Oxford.
- Keulartz, Jozef, Maartje Schermer, Michiel Korthals, and Tsjalling Swierstra. 2004. "Ethics in Technological Culture: A Programmatic Proposal for a Pragmatist Approach." *Science, Technology, & Human Values* 29 (1): 3–29. <https://doi.org/10.1177/0162243903259188>.
- Mearian, Lucas. 2015. "Insurance Company Now Offers Discounts -- If You Let It Track Your Fitbit." *Computerworld*. April 17, 2015. <https://www.computerworld.com/article/2911594/insurance-company-now-offers-discounts-if-you-let-it-track-your-fitbit.html>.
- Noorman, Merel. 2012. "Computing and Moral Responsibility." *Stanford Encyclopedia of Philosophy*.
- Richterich, A. 2018. "The Big Data Agenda: Data Ethics and Critical Data Studies." <https://doi.org/10.16997/book14.b>.
- Rip, Arie. 2013. "Pervasive Normativity and Emerging Technologies." In *Ethics on the Laboratory Floor*, edited by Simone van der Burg and Tsjalling Swierstra, 191–212. London: Palgrave Macmillan UK. [https://doi.org/10.1057/9781137002938\\_11](https://doi.org/10.1057/9781137002938_11).
- Tanner, Adam. 2014. "Different Customers, Different Prices, Thanks To Big Data." *Forbes*. 2014. <https://www.forbes.com/sites/adamtanner/2014/03/26/different-customers-different-prices-thanks-to-big-data/>.