



Vie privée, valeur des données personnelles et régulation

Grazia Cecere, Matthieu Manant

► To cite this version:

Grazia Cecere, Matthieu Manant. Vie privée, valeur des données personnelles et régulation. Annales des Mines - Enjeux Numériques, 2018, 2, pp.14 - 19. hal-01988482

HAL Id: hal-01988482

<https://hal.science/hal-01988482>

Submitted on 21 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vie privée, valeur des données personnelles et régulation

Par Grazia CECERE

IMT, TEM

et Matthieu MANANT

Université Paris Sud, RITM

Le rôle des données personnelles en économie a été accentué par la numérisation croissante de l'économie, car il est désormais possible de collecter, stocker et traiter des quantités énormes de données à des coûts de plus en plus faibles. Si l'exploitation des données personnelles permet aux entreprises de faire aux utilisateurs de meilleures offres en aidant à réduire les coûts de recherche, elle peut également être à l'origine de discriminations préjudiciables ou de sollicitations non souhaitées et inappropriées pour ces mêmes utilisateurs (Acquisti *et al.*, 2017). Cette utilisation secondaire des données personnelles peut survenir sans que l'utilisateur n'en ait conscience. D'un point de vue économique, cette exploitation possible des données personnelles est liée au fait qu'il s'agit d'un bien d'information non rival et non excluable (Cecere *et al.*, 2017). La littérature académique en économie de la vie privée s'intéresse au lien entre les comportements de divulgation de données personnelles des individus et les stratégies d'innovation des entreprises lorsqu'elles utilisent ces données. L'omniprésence de l'Internet – y compris celle de l'Internet des objets – a accru l'importance des données dans tous les secteurs, et plus particulièrement dans la publicité, ainsi que sur les sites d'e-commerce et les plateformes en ligne. Récemment, une littérature académique importante en économie et en marketing a commencé à s'intéresser aux comportements individuels dans différents contextes (Acquisti *et al.*, 2012) et, en particulier, à étudier l'efficacité de la publicité personnalisée des entreprises (Lambrecht et Tucker, 2013) et l'impact de la réglementation sur la protection de la vie privée (Campbell *et al.*, 2015).

Quelle valeur des données personnelles ?

Les entreprises de l'Internet parviennent de mieux en mieux à cibler leurs utilisateurs par une exploitation de plus en plus fine de leurs données personnelles. Si la valeur des données personnelles ainsi générée par les entreprises augmente sans cesse, la mesure ou l'estimation de cette valeur restent en pratique encore difficiles. L'analyse économique des choix et des arbitrages des individus et des entreprises sur les questions de vie privée donne une première appréciation de cette valeur. Pour les individus, le choix de partager leurs données sur les plateformes en ligne permet d'accéder à des services personnalisés et souvent gratuits (Acquisti, 2010). La personnalisation de ces services est souvent améliorée par la fréquentation de ces plateformes, les utilisateurs bénéficiant alors d'externalités liées à la divulgation de données par les autres utilisateurs. Par exemple, en faisant des recommandations – bonnes ou mauvaises – pour des produits qu'ils connaissent, les utilisateurs de sites Internet révèlent leurs préférences et ces informations peuvent ensuite bénéficier aux autres utilisateurs qui ont alors accès à cette information, soit directement, soit sous forme agrégée. La mise à profit de la divulgation de données personnelles n'est cependant pas le seul moyen pour les entreprises d'en recueillir. Les entreprises peuvent en effet s'appuyer sur les données d'activités, comme la navigation ou la localisation, pour prévoir les caractéristiques de leurs clients (Kosinski *et al.*, 2013). Les données personnelles collectées sont alors très diverses et directement personnelles, comme l'âge, l'adresse, le sexe, des préférences révélées par les médias sociaux, les achats ou des commentaires, etc. Ces données peuvent

engendrer encore plus de valeur lorsqu'elles sont combinées avec de grandes bases d'autres données (Lambrecht et Tucker, 2017 ; *The Economist*, 2010).

Les littératures académiques en organisation industrielle et en marketing étudient principalement la façon dont les entreprises exploitent les données personnelles et la manière dont cette exploitation peut créer ou stimuler de nouveaux modèles d'affaires, et donc générer de l'innovation. Les collectes de bases de données très volumineuses, l'analyse de ces données et les progrès du marketing permettent une exploitation de ces données à un niveau sans précédent et qui s'améliore régulièrement (Acquisti, 2014). La valeur des données repose donc sur la façon dont les entreprises les intègrent à leur business model. La littérature théorique repose largement sur l'hypothèse que les agents – individus et entreprises – attribuent une valeur aux données, et que les stratégies des entreprises sont affectées par les quantités et la qualité des données. Ainsi, les modèles théoriques soulignent à quel point les stratégies et modèles d'affaires des entreprises reposent sur la divulgation de données personnelles par les utilisateurs, et donc sur les préférences des consommateurs dans leur vie privée. Plus généralement, ces approches théoriques supposent une acceptation par les utilisateurs du modèle de l'Internet – services gratuits en échange de données personnelles. Fudenberg et Tirole (1998) montrent ainsi qu'un monopole vendant des biens durables peut adopter des stratégies différentes selon les types de consommateurs – anonyme, semi-anonyme ou identifié. Cependant, à notre connaissance, peu de travaux empiriques cherchent à estimer la valeur réelle des données personnelles pour les entreprises, bien que ces données soient clairement valorisées par les entreprises. Un rapport de l'OCDE (2013) suggère toutefois différentes méthodes détaillées d'évaluation des données personnelles et propose un aperçu des entreprises qui opèrent sur le marché en incluant les sociétés de l'Internet directement en lien avec les utilisateurs, mais également des entreprises tierces qui échangent et traitent ces données. En particulier, le rapport s'intéresse à la valeur des données sur les marchés des *data brokers* en montrant que la valeur des données personnelles est liée à la qualité des données collectées comme à leur volume.

Comment les firmes valorisent-elles les données ?

Pour Acquisti *et al.* (2016), trois marchés liés aux données personnelles coexistent : (1) un marché où les entreprises proposent des services aux utilisateurs en échange de leurs données personnelles, (2) un marché où les individus paient pour se protéger, et (3) un marché des *data brokers* – des sociétés tierces – où des entreprises collectent des données et les utilisent pour une commercialisation en *Business-to-Business* (B-to-B).

Sur le marché primaire qui regroupe les sociétés de l'Internet et leurs clients, les données personnelles permettent de concevoir des campagnes publicitaires plus efficaces et de fixer des prix proches de la disposition à payer des individus. Pour la plupart des entreprises en ligne, les revenus publicitaires représentent une source majeure de revenus (Martin et Murphy, 2016). L'exploitation des données de navigation, qui fournissent des informations détaillées sur la façon dont les individus interagissent avec les sites web et la publicité, ainsi que l'utilisation croissante des algorithmes contribuent dans un second temps à accélérer le rythme de l'innovation. Varian (1997) souligne ce double usage des données personnelles par les sociétés de l'Internet : la première utilisation facilite les interactions des entreprises avec leurs clients, alors que la seconde utilisation peut supposer une transmission de ces données à une ou plusieurs entreprises tierces mieux à même d'exploiter les données personnelles. L'exploitation des données personnelles permet une discrimination par les prix de premier degré ou, de façon plus réaliste, une discrimination de troisième degré ⁽¹⁾, car

(1) La discrimination du premier degré est permise par la connaissance des préférences de chaque individu, celle du deuxième degré par la connaissance des préférences de sous-groupes non identifiés, et celle du troisième degré par la connaissance des préférences de sous-groupes identifiés.

ces données permettent à une entreprise d'identifier le prix de réserve d'un individu. Dans cette perspective, le lien entre les stratégies des entreprises et les données personnelles, et donc la vie privée, apparaît central (Taylor, 2004). En suivant cette hypothèse d'un effet positif de la discrimination pour les utilisateurs, un travail théorique récent de Belleflamme et Vergote (2016) suggère que l'utilisation de technologies pour dissimuler des informations personnelles pourrait réduire le surplus du consommateur. Outre la discrimination par les prix, d'autres formes de discrimination préjudiciables aux individus peuvent cependant également exister, ce qui pose la question de l'usage fait des données personnelles et donc des conséquences sur la confiance des utilisateurs. Ainsi, des travaux récents montrent que, sur le marché du travail, les recruteurs peuvent discriminer entre candidats sur la base des informations trouvées sur les médias sociaux (Acquisti et Fong, 2015 ; Manant *et al.*, 2015). Lambrecht et Tucker (2018) ajoutent à ces travaux le rôle des algorithmes utilisés par les médias sociaux qui reproduisent les discriminations hors ligne en s'appuyant sur des bases de données elles-mêmes biaisées. Plus globalement, cette littérature académique souligne comment les entreprises peuvent exploiter les informations personnelles que les utilisateurs « laissent » en ligne.

Sur le marché secondaire qui regroupe les sociétés de l'Internet et les tiers, l'exploitation des données personnelles est réalisée à l'échelle du marché quand les données personnelles sont achetées par des sociétés tierces, telles que de nombreuses sociétés de marketing comme BlueKai et Avarto spécialisées dans la gestion de données. L'utilisation secondaire de données personnelles par des tiers survient lorsqu'elles sont transmises à des sociétés telles que des courtiers de données, des agrégateurs de données, des annonceurs ou, plus largement, à des entreprises qui ont la compétence de les exploiter (Akçura et Srinivasan, 2005). L'utilisation par des tiers et l'utilisation secondaire de données personnelles au sein d'une même entreprise semblent moins légitimes si des données personnelles sont envoyées ou utilisées à l'insu de l'utilisateur.

Quelle valeur des données pour la régulation ?

L'émergence de nouvelles entreprises dont les modèles d'affaires sont basés sur l'exploitation des données personnelles a attiré l'attention des régulateurs sur la nécessité de trouver un juste équilibre entre protection de la vie privée et promotion du partage de données pour encourager l'innovation et améliorer les services (Tucker, 2012). La définition de la source de valeur des données est indispensable pour bien définir l'impact de la régulation sur les stratégies des entreprises. L'intervention publique sur les questions de vie privée reste cependant complexe. Premièrement, l'innovation dans les secteurs où les données personnelles jouent un rôle-clé est difficile et la concurrence entre les acteurs est forte. Deuxièmement, l'élaboration d'une réglementation visant à protéger la vie privée des individus peut avoir une influence sur les choix de ces mêmes individus (Marthews et Tucker, 2017), mais également sur les choix des entreprises (Miller et Tucker, 2009). L'examen des instruments utilisés par les autorités de réglementation permet de comprendre comment la régulation sur les questions de vie privée peut avoir des effets sur les fonctionnements des marchés.

Aux États-Unis, où la Federal Trade Commission (FTC) fournit des directives au niveau sectoriel, c'est l'autorégulation qui prévaut et le principe est de stimuler la « concurrence sur la vie privée » tout en tenant compte des défaillances de marché (Cecere et Rochelandet, 2012). Cette approche considère les deux côtés du marché en supposant que les consommateurs peuvent décider de garder confidentielles leurs préférences en matière de vie privée et que les entreprises se conforment au respect du principe de transparence et de contrôle des questions de vie privée, par exemple en fournissant des informations détaillées sur la collecte des données. Dans cette perspective, les politiques de confidentialité, souvent au travers des chartes de vie privée, reposent sur un principe de notification et de consentement selon lequel les individus sont censés lire les politiques de confidentialité des entreprises et choisir de consentir, ou non, aux conditions de service (Cranor, 2012). Dans ce cadre de régulation, les politiques de confidentialité doivent fournir des informa-

tions suffisantes aux individus sur la manière dont les entreprises collectent, utilisent, partagent et sécurisent les données personnelles (Marotta-Wurgler, 2016). Cependant, des preuves empiriques montrent que ces politiques sont trop longues à lire et trop complexes à comprendre pour un non-spécialiste (McDonald et Cranor, 2008). L'approche de la FTC a ainsi encouragé la création de services de certification en ligne par des tiers privés comme TRUSTe et BBB avec leurs labels ou sceaux de vie privée qui aident à réduire, pour les utilisateurs, les coûts cognitifs liés à l'évaluation des risques. Néanmoins, le recours à ces services soulève un problème de sélection adverse. Des recherches empiriques montrent en effet que les sites Internet certifiés par TRUSTe sont plus de deux fois davantage susceptibles de ne pas être dignes de confiance que les sites non certifiés (Edelman, 2011). Ces résultats suggèrent la nécessité d'une intervention du régulateur pour assurer la qualité des sceaux privés.

En Europe, l'approche réglementaire est davantage axée sur la mise en place d'un cadre général visant à protéger la vie privée des consommateurs dans tous les secteurs, ce qui est une approche notablement différente de l'approche américaine qui s'appuie principalement sur l'autorégulation. Il s'agit dans le cas européen de limiter les effets de sélection adverse en garantissant davantage aux individus un cadre rigoureux de protection des données personnelles que doivent respecter les entreprises. Le Règlement général sur la Protection des Données (RGPD), entré en vigueur le 25 mai 2018, impose désormais aux entreprises l'obligation de demander aux consommateurs leur consentement pour utiliser leurs données, et permet aux individus d'accéder à plus d'informations sur la façon dont leurs données sont traitées par les entreprises. Il renforce également le droit des individus à l'oubli de leurs données, c'est-à-dire que ces derniers peuvent demander la suppression ou la modification de leurs données à ceux qui les détiennent. La directive européenne encourage la protection de la vie privée dès la conception (*privacy-by-design*), en favorisant la prise en compte des questions de vie privée dès les premières phases, puis tout au long du développement d'un bien ou d'un service. Ce règlement, adopté en 2016, a remplacé le précédent cadre législatif européen en mettant à jour la précédente directive 95/46/CE sur la protection des données et la directive sur la vie privée et les communications électroniques 2002/58/CE. Globalement, si l'approche européenne est censée apporter aux individus plus de transparence et de protection qu'aux États-Unis, elle devrait également être plus coûteuse pour les entreprises qui s'y conforment, car elle leur impose des obligations plus fortes.

Bien que la réglementation de la vie privée s'adresse aux consommateurs et aux entreprises, elle peut également avoir des conséquences indirectes sur la structure du marché. Dans l'ensemble, la direction et la taille de ces effets ne sont pas claires. Alors que la réglementation aide à créer un cadre clair pour les entreprises, il apparaît nécessaire de comprendre le rôle global joué par les données personnelles sur les marchés. Pour cette raison, nous nous sommes concentrés ici sur les principes régissant la réglementation de la vie privée et sur les preuves théoriques et empiriques de l'impact de cette réglementation sur les marchés, mais aussi sur la façon dont les technologies et les violations de données personnelles peuvent influencer ces marchés.

Références

- ACQUISTI A. (2010), "The Economics of Personal Data and the Economics of Privacy", OECD Conference Centre.
- ACQUISTI A. (2014), "From the economics of privacy to the economics of big data", *Privacy, Big Data, and the Public Good: Frameworks for Engagement*; Stefan Bender, Julia Lane, Helen Nissenbaum, and Victoria Stodden (eds), 76-95.
- ACQUISTI A. & M. FONG C. (2012), "An Experiment in Hiring Discrimination Via Online Social Networks", *SSRN Electronic Journal*.

- ACQUISTI A., JOHN L. K. & LOEWENSTEIN G. (2012), "The Impact of Relative Standards on the Propensity to Disclose", *Journal of Marketing Research*, 49, 160-174.
- ACQUISTI A., TAYLOR C. & WAGMAN L. (2016), "The economics of privacy", *Journal of economic Literature*, 54(2), 442-492.
- AKÇURA M. T. & SRINIVASAN K. (2005), "Research note: customer intimacy and cross-selling strategy", *Management Science*, 51(6), 1007-1012.
- BELLEFLAMME P. & VERGOTE W. (2016), "Monopoly price discrimination and privacy: The hidden cost of hiding", *Economic Letters*, 149, 141-144.
- CAMPBELL J., GOLDFARB A. & TUCKER C., "Privacy regulation and market structure", *Journal of Economics & Management Strategy* 24 (1), 47-73.
- CECERE G., LE GUEL F., MANANT M. & SOULIÉ N. (2017), "The economics of privacy", *The New Palgrave Dictionary of Economics*.
- CECERE G. & ROCHELANDET F. (2012), « Modèle d'affaires numériques, données personnelles et sites web. Une évaluation empirique », *Revue française de gestion*, 224, 111-124.
- CRANOR L. F. (2012), "Necessary but not sufficient: Standardized mechanisms for privacy notice and choice", *J. on Telecomm. and High Tech. L.*, 10, 273-307.
- EDELMAN B. (2011), "Adverse Selection in Online 'Trust' Certifications and Search Results", *Electronic Commerce Research and Applications*, 10(1), 17-25.
- FUDENBERG D. & TIROLE J. (1998), "Upgrades, tradeins, and buybacks", *RAND Journal of Economics*, 29(2), 235-258.
- KOSINSKI M., STILLWELL D. & GRAEPEL T. (2013), "Private traits and attributes are predictable from digital records of human behavior", *Proceedings of the National Academy of Sciences*, 110(15), 5802-5805.
- LAMBRECHT A., GOLDFARB A., BONATTI A., GHOSE A., GOLDSTEIN D., LEWIS R., RAO A., SAHNI N. & YAO S. (2014), "How do firms make money selling digital goods online?", *Marketing Letters*, 25, 331-341.
- LAMBRECHT A. & TUCKER C.E. (2013), "When does retargeting work? Information specificity in online advertising", *Journal of Marketing research*, 50(5), 561-576.
- LAMBRECHT A. & TUCKER C.E. (2017), "Can Big Data protect a firm from competition?", *CPI Antitrust Chronicle*.
- LAMBRECHT A. & TUCKER C.E. (2018), "Algorithmic bias? An empirical study into apparent gender-based discrimination in the display of STEM career ads", *Social Science Research Network*. WP.
- MANANT M., PAJAK S. & SOULIÉ N. (2018), "Can Social Media lead to Labour Market Discrimination: A Field Experiment", WP.
- MAROTTA-WURLER F. (2016), "Self-regulation and competition in Privacy Policies", *Journal of Legal Studies*, 45(S2), 13-39.
- MARTHEWS A. & TUCKER C. (2017), "Government Surveillance and Internet Search Behavior", SSRN.
- MARTIN K. D. & MURPHY P. E. (2016), "The role of data privacy in marketing", *Journal of the Academy of Marketing Science*, forthcoming.

- McDONALD A. M. & CRANOR L. F. (2008), "The Cost of Reading Privacy Policies", *I/S: A Journal of Law and Policy for the Information Society*, 4(3), 540-565.
- MILLER A. R. & TUCKER C. (2009), "Privacy protection and technology diffusion: The case of electronic medical records", *Management Science*, 55(7), 1077-1093.
- OCDE (2013), "Exploring the Economics of Personal Data : A Survey of Methodologies for Measuring Monetary Value ", *OECD Digital Economy Papers*, N° 220, Éditions OCDE, Paris.
- TAYLOR C. A. (2004), "Consumer privacy and the market for customer information", *RAND Journal of Economics*, 35(4), 631-665.
- The Economist* (2010), "Data, data everywhere", *The Economist Newspaper Limited*.
- TUCKER C. E. (2014), "Social Networks, Personalized Advertising, and Privacy Controls", *Journal of Marketing Research*, 51(5), 546-562.
- TUCKER C. E. (2012), "The Economics of Advertising and Privacy", *International Journal of Industrial Organization*, 30(3), pp. 326-329.
- VARIAN H. R. (1997), "Economic aspects of personal privacy", *Privacy and Self-Regulation in the Information Age*, US Department of Commerce.