

Open-Data: A Solution When Data Constitutes an Essential Facility?



Claire Borsenberger, Mathilde Hoang, and Denis Joram

1 Introduction

Data is a type of raw material, most of the time unstructured, derived from observations, experiments, measures or computations, collected by a wide range of organizations and institutions. Once it has been analyzed or interpreted thanks to intelligent methods like data mining,¹ data becomes information (a set of contextualized and structured data making them meaningful) suitable for making decisions. Thanks to digital technology, data collection and processing have become easier and less costly. Consequently, the amount of data collected has increased exponentially during the last decades, and many digital firms have built their strategy and business model on data. Some firms have developed entirely new business models directly based on data monetization, such as data brokers, specialized in data collection, that process and analyze data in order to resell information to other economic actors.

Claire Borsenberger: Senior Economist, Head of ‘Doctrine and Modelization’ Department; Direction of Institutional Affairs and Regulation—Groupe La Poste. The opinions expressed here are ours and do not necessarily reflect the position of La Poste.

Mathilde Hoang: Apprentice in the department ‘Doctrine and Modelization’; Student of Master “Network Industries and Digital Economy”, TELECOM ParisTech.

Denis Joram: Head of Regulation and Study Division; Direction of Institutional Affairs and Regulation—Groupe La Poste.

¹Data mining is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.

C. Borsenberger (✉) · D. Joram
Groupe La Poste, Paris, France
e-mail: claire.borsenberger@laposte.fr

M. Hoang
TELECOM ParisTech, Paris, France

Other firms, such as Internet platforms (Facebook, Google Search, YouTube, as examples) provide “free” services to their customers in exchange of their data that they monetize.

Thanks to appropriate data algorithms, firms, especially those on-line, are able to extract detailed knowledge about consumers and markets. This raises the question of the essential facility character of data. Moreover, the features of digital markets lead to a concentration of this core input in the hands of few big “superstars” and arouse legitimate economic and societal concerns. In a more and more data-driven society, one could ask if data openness is a solution to deal with power derived from data concentration.

Section 2 examines the essential facility character of data. Section 3 summarizes recent initiatives to extend open data policies in France and at the European Union, and analyzes the consequences of a possible extension of openness to privately held data in light of economic theory with a particular focus on the case of postal operators. Section 4 concludes.

2 Is Data an Essential Facility that Should Be Opened?

2.1 *Data Is an Impure Public Good*

In economists’ language, data is a non-rival good (Isaac, 2016; Lambrecht & Tucker, 2015; Sokol & Comerford, 2017), meaning that one person’s consumption does not preclude another’s: the collection and use of a piece of data by one firm does not induce its disappearance (contrary to the consumption, for instance, of an apple—a “private good”).

Opinions are less definite over the non-excludable character of data, i.e. the absence of gateways on consumption.² According to Sokol and Comerford, no firm can, or does, control all of the world’s data. It is difficult to request financial compensation to access a piece of data (Lambrecht & Tucker, 2015). If one provider has a piece of data, another provider is not prevented from collecting that very same piece of data.

However, companies are becoming increasingly aware that they are sitting on huge amounts of under-utilized data and looking for ways to increase its value. The fact that data is a strategic asset—data can improve decision making, produce more valuable goods and help optimize production—that has commercial value could be a barrier to disclose it freely, leading to situations in which few actors control it, excluding some potential users from its “consumption”.

Excluding some users from essential data raises concerns about competition. New product or service providers could be prevented from entering a market.

²For example, fresh air is non-excludable, because it is impossible to stop several people in the same area from breathing the same fresh air.

Even established providers could be forced to exit a market if they do not have access to some “essential” data. These threaten the viability of a competitive environment. In this context, two questions arise: First, is a given type of data an essential facility? Second, if yes, is open data the right solution to guarantee access to this essential facility?

2.2 *Is Data an Essential Facility?*

Heitzler (2009, p. 80) defines essential facilities as “inputs that are unconditionally necessary to provide certain goods or services and that are unfeasible or too costly to be duplicated or to be bypassed. At the same time there must not exist sufficient demand-side substitution possibilities for the service itself. Shortly, essential facilities not only have to be nonreplicable but also non-substitutable with regard to the service they are needed for”. For example, data is sometimes considered as the new oil that drives the economy; without it, progress would halt. According to Gans (2018), like internal combustion engines with oil, data-driven markets need data to run. Digital firms rely on data to offer their services. For instance, Spotify’s recommendation algorithm relies on past behavior of users to improve its recommendations. In this context, one could argue that new entrant who, by definition, does not own a database on users’ profile and usages similar to incumbents, cannot compete with the latter. It seems then necessary to give entrants access to incumbents’ data.

Isaac (2018) does not share this point of view. By itself, raw or primary data has no meaning: the simple fact to collect data is not sufficient to create profit-enhancing opportunities. The economic value of primary data comes from its consolidation with metadata,³ its treatment by algorithms that transforms primary data into knowledge by crossing it with other data. This process of transformation and value creation lies on technical investments (collection and treatment infrastructure) but above all on human, organizational and strategic capabilities (Isaac, 2016). This view is shared by Lambrecht and Tucker (2015, p. 11), who claimed, “It is only when combined with managerial, engineering, and analytic skill in determining the experiment or algorithm to apply to such data that it proves valuable to firms”.

Furthermore, Lambrecht and Tucker (2015) argued that data held by incumbents cannot be defined as non-replicable or rare. First, as we saw, it is a non-rival good with a near-zero marginal cost of reproduction. Second, tools and technologies to

³Metadata is data (information) that provides information about other data. For example, a digital image may include metadata that describes how large the picture is, the color depth, the image resolution, when the image was created, the shutter speed, and other data.

collect, gather, store and analyze data are more and more powerful and affordable. Lambrecht and Tucker (2015) and Rubens (2014) speculate that storage costs may eventually approach zero; Altman, Nagle, and Tushman (2015) argued that information costs are rapidly approaching zero. Third, firms have developed a business model based on the sale of databases. Fourth, consumers leave more and more traces of their needs and preferences, sometimes unconsciously across the Internet. Moreover, entry into some digital markets, such as social networks, is also facilitated by the fact that consumers are not reluctant to use different services if the opportunity cost to multi-home is not high. Finally yet importantly, the value of some data decreases through time (Sokol & Comerford, 2017). In this case, the main concern of entrants should not be to get the incumbent's data but to collect updated and differentiated data to respond to sometimes evolving needs of users (Schepp & Wambach, 2015).

These arguments suggest that the market power of dominant firms (in other words, the roots of the dominant position of some "superstars" or giant tech) comes more from their ability to provide a reliable and high quality good, reinforced by network effects and the switching costs incurred by customers, than from primary data. However, it seems difficult to deny that to some extent the ability to satisfy consumers' needs and to exploit network effects come from information and knowledge provided by data. Moreover, one can find a counter-example that leads to the conclusion that data are a necessary or essential resource in the digital economy. Indeed, in some cases, the collection or reproduction of data will be costly, consumer will be reluctant to multi-home, the quality of data offered by third parties will be lower, and so on (see for instance Autorité de la concurrence and Bundeskartellamt, 2016; Graef, 2016; Grunes & Stucke, 2015).

Academic literature remains divided on whether data is an essential facility. The answer cannot be unequivocal: it depends on the type of data and market under review. For instance, in the markets on which Graef (2016) and Grunes and Stucke (2015) rely to support their opinions, search engines (the Google Search case) and digital maps (the TomTom/Tele Atlas case) collect a huge amount of data is an essential pre-requisite to develop their services. Such large datasets could be considered as an entry barrier since the costs of collecting, processing and storing the data are generally high whereas it has (almost) zero marginal cost of reproduction.

On the contrary, for social networks, which nevertheless exhibits high network effects and switching costs based to some extent on users' data, access to data has not protected incumbents from competition. This industry has experienced a succession of large firms: MySpace replaced Friendster and then was replaced by Facebook as the leading social network site. Facebook could be in the future replaced by another actor such as Instagram or a not yet existing actor (Lambrecht & Tucker, 2015). It seems more appropriate to have a case-by-case approach rather than to establish *per se* rules.

3 Towards a More General Open (Public and Private) Data Environment and a “Common European Data Space”?

For at least 10 years, we have observed a move from a closed proprietary data resources to a common shared resource, notably under the impetus of “Open Government Data” (OGD) policies. “Open data” is a piece of data or content that anyone is free to access, use, reuse, and redistribute (European Commission, 2014). This means that the piece of data is available in a convenient and modifiable form and under terms that permit its reuse, redistribution and mixing with other datasets by everyone (open license). There should be no discrimination against fields of endeavor or against persons or groups. For example, restrictions that would prevent ‘commercial’ use, or restrictions of use to certain purposes (e.g. only for research or in education) are not allowed. Accessibility implies affordability. Such data must be available at no more than a “reasonable” reproduction cost to be qualified as “open”.

“Open government data” or open “public sector information” are defined as open data or information generated, created, collected, processed, preserved, maintained, disseminated or funded by or for the Government or public institutions (OECD, 2006). Open data policies have usually materialized with public sector datasets becoming easily accessible and reusable by the general public through governmental web portals. The Obama Administration was the first to launch an open governmental data portal (data.gov) in May 2009, rapidly followed by countries around the world (e.g., UK, Spain, Singapore, Australia, Chile, France). The original focus was on governmental data; recent initiatives extend obligations of openness to data held by private actors.

3.1 *The French Precedent and the Revised PSI Directive at the European Level*

In France, the law for a digital republic that came into force on October 7, 2016 has already introduced new provisions in French law to bolster and broaden the open data policy. Article 3(I) obliges not only central and local governments, but also public and *private legal entities having a public service mandate*, to exchange public information they produce or receive. Article 4 of the bill sets up a new public service mandate under which the government is tasked with making available and disseminating a new class of public data named “*benchmark data*”⁴ (or *high-value data*) to foster its re-use. Last but not least, the Act introduced the new concept of “data of general interest” by expanding the open data policy to include public and

⁴For the moment, nine datasets have been identified as benchmark data: National Address Database, Enterprise Database (SIRENE file), Geographic official code, Cadastral Map, Landing register, Reference document on the organization of State, Big Level Reference document, National file of associations and the Reference document of jobs and professions.

private entities, public service concession holders or entities *whose activities are subsidized by the public authorities*, and by providing streamlined access for INSEE (National Institute of Statistics and Economic Studies) to some *private* databases for the purposes of mandatory statistical surveys. Article 5 introduced an obligation for public service concession holders (potentially private firms) to allow the concession-granting authority to publish, as open data, the main data concerning the activity covered by the public service concession.

At the European level, the Commission published April 25, 2018 a proposal to revise the Public Sector Information Directive (PSI) Directive (European Commission, 2018a, 2018b). During the preparation phase of Directive's review, the concept of "reverse PSI" that would entail access for public sector bodies to re-use privately held data was examined. Fortunately, reverse PSI does not appear in the published proposal. It would have raised a number of difficult questions: How to deal simultaneously with proprietary and commercially sensitive/confidential information? How to balance commercial interests with public interest? How to reconcile the need for data protection with wider access to data? How to make datasets collection sustainable if collector bodies cannot charge anymore for their reuse? Last and not least, how to distribute the costs of facilitating access to and re-use of these data for the public good (the provider, the user/consumer or the citizen) knowing that the openness of these data will inevitably need additional infrastructure, data protection and security measures, as well as better data readability and interoperability?

Nevertheless, the Commission has not completely given up the possibility of reverse PSI. As stated in the Proposal for a Directive on the re-use of public sector information published the 25th April 2018, "the scope of application of the Directive shall be extended to documents held by public undertakings active in the areas defined in the Directive 2014/25/EU on procurement by entities operating in the water, energy, transport and postal services sectors and by public undertakings acting as public service operators under Regulation (EC) No 1370/2007 insofar as they are produced as part of the provision of services in the general interest, as defined by law or other binding rules in the Member State" (p. 9).

3.2 Motivations and Expected Benefits of OGD

Two main categories of objectives or benefits are expected to be reached by opening public data. The first one is rooted in the ethos of democracy and freedom of information. Government data openness is considered a means to promote democracy, to give citizens access to information, to increase transparency of government actions and to increase the participation, interaction, self-empowerment and social inclusion of open data users (e.g. citizens) and providers (Bertot et al., 2012; Janssen, 2011).

The second main objective pursued by public institutions through open data initiatives is rooted in increasing economic value and efficiency by reducing transaction costs. Data are no longer collected several times and saved in multiple

repositories. Exchanges are simplified by promoting machine-readable interoperable formats. More transparency can reduce asymmetries of information between economic agents' which can produce principal-agent problems such as moral hazard situations—those where the more informed party makes decisions in his own profit while the cost falls on others. Lowering transaction costs will help both the public and private sectors to provide better services, to develop new production methods and to introduce new products and services, generating economic value (Jetzek, 2013).

The expected benefits of OGD have been largely outlined by a number of ex ante studies (see Carrara, Chan, Fischer, & van Steenberg, 2015 for a survey). However, ex post evaluations of concrete impacts of OGD are lacking. Koski (2015, p. 25) admits that “to [his] best knowledge, there is no reported comprehensive country-level ex-post impact assessment of opening up government data. The current research-based knowledge concerning the impacts of open data is limited only to narrow areas and largely based on case examples.” He explains this lack of ex post analysis by newness of the OGD phenomenon, the lack of systematically collected statistical data on the use of data, and the absence of indicators or models to assess the impacts of open data. According to Zuiderwijk and Janssen (2014), suitable evaluative indicators for the assessment of the success of open data policies lack. Performance indicators often concentrate on the input of policies, such as the number of datasets that are publicly available (Bertot, McDermott, & Smith, 2012). Less attention has been given to the original intent or goals of open data policies and to the reuse of data by companies and citizens.

According to Janssen, Charalabidis, and Zuiderwijk (2012), many public organizations have jumped on the bandwagon of making data available without having a sound policy. This has resulted in central portals with poor quality data that were already publicly available, without feedback mechanisms. According to the 4th edition of the Open Data Barometer released by the World Wide Web Foundation (2017), even if 79 out of the 115 governments surveyed have an open government data portal, only seven governments include a statement on open data by default in their current policies. Moreover, according to this report, the data released are usually incomplete, out of date, of low quality, fragmented, and published with no metadata or guidance documentation, which makes the data hard to use. In addition, complete datasets are often published by other government agencies or national statistics offices (NSOs) on their own platform, reducing the expected cost savings.

The impact assessment accompanying the proposal to recast the PSI directive gives a more positive view of open data policy. According to the summary report of the “high level round-table discussion on Public Sector Information re-use under the PSI Directive” held on March 16, 2018 (European Commission, 2018c), the PSI Directive has induced an increase in the supply of data as well as an increase in the demand, while ensuring fair and proportionate conditions for reuse. Deloitte (2017) stated that the Directive improved the efficiency of the public sector itself and created economic gains for public sector bodies. Furthermore, Deloitte's economic analysis showed that the Directive enabled the creation of more than 8,000 data-related jobs since 2013. The authors concluded the analysis shows that the

benefits of the PSI Directive exceeded its costs. This explains why the Commission wishes to extend the concept of open-data to other entities, in particular to public undertakings providing utilities (even if she gives up for the moment the idea of a reverse-PSI). But such an initiative neglects the risks and costs of forcing economic actors to disclose their data.

3.3 *The Risks and Costs of a Larger Open-Data Policy*

Forcing private firms to disclose their data could be counterproductive. Such a policy may destabilize and distort the economy. In particular, if a “free of charge” scheme is imposed,⁵ it could not only lead to underinvestment in data production but also harm the provision of public services when they are provided by private entities whose business models are driven by data. Such business model may become no longer sustainable and lead to the firm’s collapse.

Such a policy may also distort competition between companies forced to share their data for a reuse for free and those that could reuse these data without supporting the cost of production of this input. The classical free-riding and prisoner’s dilemma problems related to innovation, for instance, would appear.

Indiscriminately disclosing all data could also threaten individuals’ privacy and national security. In general, national laws prevent the publication of personal data that can be traced back to the individual. Despite these legal provisions aiming to protect individuals’ privacy, recent scandals show that security system can fail. (Consider for instance the Facebook/Cambridge Analytica scandal.) The most optimistic people think that the General Data Protection Regulation (GDPR), entered into force on May 25, 2018 in all European Member States, will be enough to protect privacy. But several authors underline the relative ease of re-identifying people thanks to large-scale metadata datasets. For instance, de Montjoye, Hidalgo, Verleysen, and Blondel (2013), de Montjoye, Radaelli, Singh, and Pentland (2015) showed that four spatio-temporal points are enough to uniquely identify 95% of people in a mobile phone database of 1.5 million people and to identify 90% of people in a credit card database of 1 million people. They furthermore showed that, in both cases, even coarse or blurred datasets provide little anonymity.

⁵Economists generally support the marginal cost pricing of government data (de Vries et al., 2011, Pollock, 2008), or even zero pricing (Newbery, Bently, & Pollock, 2008). The imposition of access charges above the marginal costs of producing and distributing information results in a double of burden of economic inefficiency. The first-order effect is the curtailment of the use of the information or the increased cost of using it to produce conventional commodities and services, and hence the loss of utility derived from such products by consumers. A second round of inefficiency is incurred by the inhibition of further research which otherwise would be the source of more public goods in the form of new knowledge. But according to Pénin (2013), “gratuity is not a necessary condition for achieving openness. A piece of knowledge can be considered as open even if one has to pay in order to reuse it, provided that the fees are not prohibitive” (p. 134).

Last but not least, data of only good quality should be publicized. Open access to data that is unreliable, of low quality or that provides only one point of view of a more complex issue can result in discussions, confusions, a biased picture of the situation and wrong conclusions, wastes resources and, at the end of the day, could be detrimental for transparency and even trust in the government.

3.4 The Special Case of Postal Data

As mentioned in Sect. 2, the Commission proposes to add into the scope of the PSI Directive data held by operators acting as public service operators, such as postal universal service providers. In general, a limited set of obligations will apply to those public undertakings: they can charge above marginal costs for dissemination and are under no obligation to release data they do not want to release.⁶ However, the Commission proposes to create a new category of data—data of high-value⁷—that will be defined by a delegated act. According to Article 13, these high-value datasets will have to be machine-readable, accessible via application programming interfaces (APIs), and provided for free except if an impact assessment has demonstrated that making the datasets available for free will lead to a considerable distortion of competition.

These new provisions of the PSI Directive could seriously hurt postal operators charged with universal service. Many datasets owned by postal operators risk being designated as “high-value” where allowing them to charge for their datasets, at least to recover costs of collect and maintenance, may be subject to a claim that it would lead to a “considerable” distortion of competition. How will the impact assessment be conducted? How will a “considerable” distortion of competition be determined?

Such a qualification of postal datasets could create an important distortion of competition between public undertakings and private companies that are not under the scope of the PSI directive but operate on the same markets. It could furthermore undermine the current efforts of postal operators to diversify their revenue sources by monetizing their datasets. Indeed, data monetization creates opportunities for operators that have significant data volume to leverage untapped or under-tapped information and to create new sources of revenue. Relevant data include postal codes, the name of street, the complete postal address of millions of households and businesses, the list of postal access points, and so on.

⁶According to recital 22 and article 3 of this proposal, all documents that the public undertaking are made available for re-use will fall within the scope of the Directive and will be re-usable for commercial and non-commercial purposes under the conditions set in the Directive.

⁷According to article 2, high-value datasets means documents the re-use of which is associated with important socio-economic benefits, notably because of their suitability for the creation of value-added services and applications, and the number of potential beneficiaries of the value-added services and applications based on these datasets.

Clearly, under the current recast of the PSI Directive, this source of revenue is under threat. In particular, the national address database (NAD) is generally considered as highly valuable for the society since a broad variety of services depend on accurate, up-to-date address data, including emergency services, the police, transport services, and GPS systems. Yet, today, many postal operators monetize their own NAD (for instance, Deutsche Post has created a subsidiary Deutsche Post Direkt; Royal Mail monetizes through a license system its Postcode Address File (PAF); and so on). Open data policy puts NAD revenue at risk.

Facing the open data movement, some postal operators decided themselves to open some of their databases for free. As a test to start an exchange with the open data community, Swiss Post published at the end of 2017 an initial set of non-personal information—names of locations, municipalities and streets, details of physical access points or postcode directories—on its own platform set up especially for this purpose: swisspost.ch/open-data. In France, La Poste contributed to the creation of an open free National Address Database in 2015. This French NDA is the product of a collaboration between public authorities (Etalab, a mission of the General Secretariat for the Modernization of Public Action), public actors (the National Institute of Geographic Information and Forest and La Poste Group) and civil society (OpenStreetMap).

Nevertheless, one can question the relevance of a larger open data policy that would force postal operators to disclose data for free, whereas the maintenance of such database is clearly costly. Royal Mail's costs to manage the PAF are estimated to £24.5 million per year. Most postal databases are sensitive and constitute a strategic asset. Many data collected by postal operators during their activities are the property of their clients and cannot be disclosed without their explicit consent. Furthermore, in a time of declining volumes, revenues derived from data monetization could help finance the universal service.

As stated by the European Centre of Employers and Enterprises providing public services and services of general interest (2018), “public services’ enterprises must deliver their services in a cost-efficient way. As such, they should not be forced to give out value for free or at marginal costs to other enterprises. The risk is that the EC proposal about future delegated acts forces public undertakings to make high-value datasets available for free. This would hinder ongoing innovation in public services’ enterprises by creating legal uncertainty and making investments in own data sets and existing cooperation with start-ups unstable and risky.” All these considerations, together with those already mentioned about incentives to collect data and innovate, should be taken into account by public authorities in the open data debate.

4 Conclusion

Many persuasive reasons suggest not making public all government or privately held data. The most obvious is the protection of citizens’ personal data and the safeguarding of strategic assets for private companies. Attempts to legally force

access to private data, even that classified as being of public interest, could be misleading and detrimental. Such measures could discourage market entry, investments and innovations, and thereby jeopardize the development of a future flourishing European Data Economy. Mandatory open data policy could be justified only by the existence of market failure, that is to say when private data of public interest are subject to under-provision due to antitrust issues or coordination failures. In this case, mandatory access might be a conceivable remedy.

Consequently, only a case-by-case approach should be followed to determine if obligatory access is the best solution among all other feasible remedies. Mandatory access should be used only to restore the functioning of markets, and only if it proves being the most effective and least invasive remedy. This assessment should take into account the resources needed, the competitive distortion involved by an asymmetric obligation, and the potential risks of misuse of these data against the hypothetical value that can be gained from publicizing the data. In other words, the decision should be taken on an ex ante cost-benefit analysis of disclosure.

References

- Altman, E. J., Nagle, F., & Tushman, M. L. (2015). Innovating without information constraints: Organizations, communities, and innovation when information costs approach zero. In C. Shalley, M. Hitt, & J. Zhou (Eds.), *Oxford handbook of creativity, innovation, and entrepreneurship: multilevel linkages*. Oxford: Oxford University Press.
- Autorité de la concurrence and Bundeskartellamt. (2016). *Competition law and data*. May, 10th.
- Bertot, J. C., McDermott, P. & Smith, T. (2012). Measurement of Open Government: Metrics and Process, 45th Hawaii International Conference on System Science (HICSS), pp. 2491–2499.
- Carrara, W., Chan, W. S., Fischer, S., & van Steenberg, E. (2015). *Creating value through open data: Study on the impact of re-use of public data resources*, November.
- de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Nature, Scientific Reports*, 3, 1376.
- de Montjoye, Y.-A., Radaelli, L., Singh, V. K., & Pentland, A. S. (2015). Unique in the shopping mall: On the re-identifiability of credit card metadata. *Science*, 347(6221), 536–539.
- de Vries, M., Kapff, L., Achiaga, M. N., Wauters, P., Osimo, D., Foley, P., et al. (2011). *Pricing of public sector information study*. Brussels: European Commission, Information Society and Media Directorate General.
- Deloitte. (2017). *Study to support the review of Directive 2003/98/EC on the re-use of public sector information*, FINAL REPORT, A study prepared for the European Commission DG Communications Networks, Content & Technology, SMART number 2017/0061.
- European Centre of Employers and Enterprises providing public services and services of general interest. (2018). *European Commission proposal for a review of the PSI directive risks hindering innovation and investments in public services*, Press Release, April 26th, 2018.
- European Commission. (2014). *Open Government Data & the PSI Directive*. https://www.europeandataportal.eu/sites/default/files/training_1-1_open_government-and-the-psi_en.pdf
- European Commission. (2018a). *Towards a common European data space*, COM (2018) 232 final, 25th April 2018.
- European Commission. (2018b). *Guidance on sharing private sector data in the European data economy*, SWD (2018) 125 final, 25th April 2018.
- European Commission. (2018c). *Final workshop on the study supporting the evaluation and impact assessment of the Public Sector Information re-use (PSI) Directive – Summary of the*

- discussion. <https://ec.europa.eu/digital-single-market/en/news/summary-report-final-workshop-study-supporting-evaluation-and-impact-assessment-public-sector>
- Gans, J. (2018). *Is your data really oil?*, January 17th, <https://digitopoly.org/2018/01/17/is-your-data-really-oil/>
- Graef, I. (2016). *Data as essential facility, competition and innovation on online platforms*, Research Unit KU Leuven Centre for IT & IP Law (CiTiP), June.
- Grunes, A. P., & Stucke, M. E. (2015). *No mistake about it: The important role of antitrust in the era of big data*. The Antitrust Source, April.
- Heitzler, S. (2009, March). Traditional regulatory approaches and the postal service market, competition and regulation in network industries. *Intersentia*, 10(1), 77–106.
- Isaac, H. (2016). *Données, valeur et business model*, Les Cahiers Scientifiques de la Chaire IESO, 2016, n°21.
- Isaac, H. (2018). Digital Data: Public Good or Source of Profit. *Pouvoirs, revue française d'études constitutionnelles et politiques*, n° 164, La Datacratie, pp. 75–86.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446–456.
- Janssen, M., Y. Charalabidis and A. Zuiderwijk (2012), “Benefits, adoption barriers and myths of open data and open government”, *Information Systems Management*, vol. 29, n 4, pp. 258–268.
- Jetzek, T. (2013). *The value of Open Government Data*, Perspektiv, NR. 23-2013, pp. 47–56.
- Koski, H. (2015). *The impact of Open Data – A preliminary study*. Ministry of Finance publications 15b/2015, March.
- Lambrecht, A., & Tucker, C. E. (2015). *Can big data protect a firm from competition*, December.
- Newbery, D., Bently, L., & Pollock, R. (2008). *Models of public sector information provision via trading funds*, Mimeo, February.
- OECD. (2006). *Digital broadband content: Public sector information and content*. Paris: OECD Publications.
- Pénin, Julien (2013), Are you open? An investigation of the concept of openness for knowledge and innovation, *Revue économique*, vol. 64, n 4, pp. 133–148.
- Pollock, R. (2008). *Economics of PSI*. Cambridge: University of Cambridge.
- Rubens. (2014). *Can Cloud Storage Costs Fall to Zero?*, August 5th, <http://www.enterprisestorageforum.com/storage-management/can-cloud-storage-costs-fall-to-zero-1.html>.
- Schepp, N.-P., & Wambach, A. (2015). On big data and its relevance for market power assessment. *Journal of European Competition Law & Practice*, 7, 120–124.
- Sokol, D., & Comerford, R. (2017). Does antitrust have a role to play in regulating big data? In R. Blair & D. Sokol (Eds.), *The Cambridge Handbook of antitrust, intellectual property, and high tech* (pp. 293–316). Cambridge: Cambridge University Press.
- The World Wide Web Foundation. (2017). *Open data barometer* (4th ed.), Global report, May.
- World Wide Web Foundation. (2017). *Open data barometer* (4th ed.).
- Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31, 17–29.