LE GROUPE LA POSTE

SHAPING COMPETITION POLICY IN THE ERA OF DIGITISATION

Groupe La Poste's Contribution to Panel 1's discussions:

COMPETITION, DATA, PRIVACY AND AI

Foreword

As an actor both involved and affected by the digitisation of the economy, Le Groupe La Poste welcomes the European Commission's initiatives to reflect about the implications of this evolution for competition policy. La Poste is eager to contribute to the on-going reflections to shape competition policy in the era of digitisation and more specifically to deal with issues related to competition, data, privacy and artificial intelligence. In preliminary, let's go back to the reasons why a company such as La Poste is interested in issues related to data, privacy and AI.

Over the past, La Poste, as all postal operators (POs), has provided a secure, universally accessible platform for physical commerce and communications. Since always, POs have been trusted intermediaries providing secured communications services: protection of personal information has been in the "DNA" of POs for centuries. Today, as digitalization is for example becoming common in administrative procedures, the opportunity exists to extend their trusted intermediary role into the digital age and to act as a "bridge" to facilitate the advancement of access to the digital world thanks to their large physical network of outlets which could become a place where low digital skilled people learn to use, search and communicate with digital tools. In parallel, many postal operators are developing digital services and solutions for consumers (like hybrid mail, e-letters, e-registered mail, digital mailboxes, online payment solutions, authentication and secured archiving services, personal data store). La Poste has also launched in 2015 a "platform of connected objects" on which companies can develop services around their connected objects (data storage and analysis, billing, link with physical services provided by postmen at the door, etc.) and consumers can retrieve and control the data from their personal connected objects. Artificial intelligence is also a big challenge for postal operators, given for instance its applications to autonomous (delivery) vehicles, intelligent mail boxes and so on.

All these new services are based on data, most often personal data. This puts threat on privacy as never before and raises questions about the role of data in competition on more and more data-driven markets, notably about the character (or not) of essential facility of data. The following pages will give some insights on all these issues, based on papers published by La Poste's collaborators.

Data and competition law: what do we speak about?

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. Traditionally, data is a (secondary) by-product used as an input (an information) to optimize production, improve quality, and help taking suitable decisions. Today, 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. At current pace, 2.5 quintillion bytes of data are created each day.

As noted by Borsenberger et al. (2016), "this ocean of data and shared information should increase welfare, lower search costs, boost economic productivity, reduce economic inefficiencies and improve our own experiences thanks to data analysis and the development of predictive algorithms. It could also be a source of losses, economic inequalities, and power imbalances between those who generate (more or less consciously) the data and those who control and exploit these data. (…) This raises the issue of determining the right balance between the additional societal values generated by information

disclosure and uses of data on one hand, such as public health, national security and law enforcement, environmental protection, and economic efficiency and the potential risks to individual autonomy and privacy (discrimination, exclusion, loss of control on data disclosure) on the other hand."

We consider that competition law has a role to play to reach this right balance. Indeed, in a world where customization is the rule, its character of strategic asset is more and more prominent: once data has been analysed thanks to algorithms and intelligent methods like data mining[1], firms are able to extract detailed knowledge about consumers and markets which could potentially give them a competitive advantage. Aware of the value of data, some firms have developed entirely new business models based on data monetization, such as data brokers, specialized in data collection, that process and analyse data in order to resell information to other economic actors. Last but not least, in the digital world, data is increasingly used as a sort of currency[2]: some companies, like Internet platforms (Facebook, Google Search, YouTube for instance) provide "free" services to their customers in exchange of their (personal) data that they monetize. Consumers' data can in these cases be assimilated to the "price" paid for a seemingly free service.

From competition law perspective, all these features raise several issues, in particular the questions of the essential facility character of data and the capacity and willingness for firms to use data in a an anti-competitive way, as a tool to reduce competitive intensity, to deter entry in markets or to implement anti-competitive practices. We will focus on these issues in the following pages.

Could data be considered as an essential facility?[3]

Due to the growing importance of data in the economy, some experts consider that data is the new oil and has a character of essential facility. Everyone agrees data is a non-rival good (Sokol and Comerford, 2017; Isaac, 2016; Lambrecht and Tucker, 2015), 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 a "private good"). This makes data distinct from oil, a typical private good.

However, opinions are less definite over the non-excludable character of data, i.e. the absence of gateways on consumption[4]. For instance, Lambrecht and Tucker (2015) argued that data held by incumbents cannot be defined as non-replicable or rare. First, as we just said, it is a non-rival good with a near-zero marginal cost of reproduction. Second, tools and technologies to collect, gather, store and analyse data are more and more powerful and affordable. Lambrecht and Tucker (2015) and Rubens (2014) speculate that storage costs may eventually approach zero and Altman et al. (2015) argued that information costs are rapidly approaching zero. Third, some firms have developed a business model based on the sale of databases, contributing to their disclosure. 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[5]. Finally yet importantly, the value of some data decreases through time (Sokol and 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 evolving needs of users (Schepp and Wamback, 2015). These

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

[2] With the caveat that the same set of data could be monetized several times whereas a given amount of money cannot be used multiple times.

[3] The part is extracted from Borsenberger et al. (2019) paper.

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

[5] 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 according to Lambrecht and Tucker (2015).

arguments suggest that the market power of dominant firms (in other words, the roots of the dominant position of some "superstars" or tech giants) 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 comes from information and knowledge provided by data. Moreover, it is undeniable that in some cases, first-mover advantage can be significant, that large data sets may be very difficult to replicate, that some powerful feedback and network effects are at work: firms' algorithms become more and more powerful and performant the more data they get. A company that has a lot of users, will obtain a lot of data which will feed an algorithm that will allow to attract additional users which provide additional data and feed an ever-performant algorithm. In this context, one can find counter-examples leading to the conclusion that data is a necessary or essential resource in the digital economy. For instance, in search engine or digital map markets, the collection of a huge amount of data is an essential pre-requisite to develop this type of service (Graef, 2016; Grunes and Stucke, 2015). Furthermore, in some cases, consumer may be reluctant to multi-home, the quality of data offered by third parties may be lower, and so on (Autorité de la concurrence and Bundeskartellamt, 2016), giving to primary data a greater value.

Considering all these arguments, we think that the character of essential facility of a set of data or not depends on the type of data and market under review. In this context, competition authorities should have a case-by-case approach to determine if some data should be considered as an essential facility or not and if companies have implemented anti-competitive strategies to prevent rivals to have access to some data in order to protect or reinforce their market share, rather than establish *per se* rules.

Could firms implement anti-competitive practices based on data exploitation? Could accumulation of data lead to abuse of dominance?

Having a dominant position is not problematic from the competition law point of view but abusing from it, is. In the same way, the accumulation of data is not, in itself, problematic from the competition law point of view. However, exploiting a dataset to restrict competition or to abusively reinforce dominant position in some markets raises concerns.

The difficulty for competition authorities is to draw the line between a pro-competitive exploitation of data and feedback effects (more data that allows you to make better product, to attract more users and to collect more data) and anti-competitive behaviour (when companies use their market power to restrict competitors' access to data or use their data to enter entry, to force competitors to exit the market, or to alter the competitive equilibrium for its own benefit).

For instance, more and more discussions are arising about the possibility of tacit price collusion induced by AI (see for instance Ezrachi and Stucke, 2015). Algorithmic collusion arises when separate algorithms used by rivals, decide by themselves that the best way to maximise their respective company's profit, consists in collaborating on price. Today, opinions diverge about the reality of such coordinated mechanisms. But if this threat is credible, it poses new challenges to antitrust authorities: could these coordination situations, without any human intervention, be qualified as cartel or anti-competitive agreements? Where would the liability lie? How could such cartel behaviour be detected and proved?

In the same way, the use of 'big data' in order to engage in first-degree price discrimination, charging each consumer a different price for the same good or service, has been denounced several time. However, most economists recognize that price differentiation justified by objective differences between consumers' profile, could be beneficial from both an efficiency and an equity point of view.

On the contrary, the risk that a platform like Amazon (who operates both as a e-seller and as an intermediary platform connecting e-sellers and e-buyers) gathers data on affiliated businesses' activity

in order to extort value from those businesses, to promote its own products and services, or to thwart nascent competitors in ancillary lines of business, exists and such a practice would be clearly anti-competitive. This is confirmed by the fact that the European Competition Commissioner Margrethe Vestager announced on 19th September that her office is in the early stages of gathering on Amazon's use of their data. The issue, she said, is whether Amazon is using data from the merchants it hosts on its site to secure an advantage in selling products against those same retailers.

Another obvious form of anti-competitive practices involving data would be an agreement for exclusive licensing or exclusive access to an essential and non-replicable data set, refusing giving access to these data to all or some rivals.

A rather "new" anti-competitive agreement would be an agreement aiming to reduce competition based on privacy protection. Today, as consumers are more and more sensible to the protection of their personal data, firms could compete not only on price and quality of the service they provide but also on the guarantee they offer regarding privacy protection (which becomes part of quality of service and a parameter of non-price competition). As company colluded on prices in the "old" world, one could imagine that firms collude on privacy policy and conclude agreements to reduce competition on this aspect. Rivals may agree on a lower level of protection compared to the competitive one, minimizing their costs and reducing competitive intensity on this feature. This potential new type of anti-competitive practices pleads for not leaving privacy issues aside to competition law.

Is open data policy a solution to remedy to competitive concerns?

For its proponents, in a more and more data-driven society, data openness is the solution to deal with economic and societal concerns linked to the concentration of data in the hands of few big superstars. According to us, openness could be a solution under some particular circumstances but must not be set up as a general principle.

For at least ten 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. If the original focus was on governmental data, recent initiatives aim to extend obligations of openness to data held by private actors. In France, for instance, the law for a digital republic that came into force on October 7, 2016 bolsters and broadens the open data policy. The law 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, introduce the concept of "data of general interest" and create a new class of public data named "*benchmark data*"[6] (or *high-value data*).

At the European level, during the preparation phase of the Public Sector Information (PSI) Directive's review, the concept of "reverse PSI" that would entail access for public sector bodies to re-use privately held data was examined (European Commission, 2018a; 2018b). Fortunately, reverse PSI does not appear in the proposal published the 25th April 2018.

According to us, 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, it could not only lead to underinvestment in data production but also harm the provision of public services when they are provided by *private legal entities having a public service mandate*. In particular, this could be the case of operators in charge of SGEI like the Group La Poste. Many datasets owned by such operators risk being designated as "high-value" [7]. Such a qualification of "high value" datasets could

[6] For 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 if associations and the Reference document of job and professions.

[7] According 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.

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. In the postal case, 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.

Last but not least, 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 wrongdoing show that security system can fail. Consider for instance the Facebook/Cambridge Analytica wrongdoing. 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 et al. (2013; 2015) showed that 4 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.

Before concluding, we would like to emphasize that not disclosing all data does not mean banning all data sharing. As shown by Borsenberger et al. (2016), a general ban on the sharing of personal data would work for the detriment of all parties.

Conclusion

Many persuasive reasons suggest neither making all data public, nor imposing a general ban on data sharing are good solutions to promote competition and fight against anti-competitive behaviours in data-driven sectors. Such measures could discourage market entry, investments and innovations, and thereby jeopardize the development of a future flourishing European Data Economy and have an overall negative impact on social welfare.

Consequently, only a case-by-case approach should be followed by antitrust authorities when they examine the impact of data on competition and decide the right solution of prevent anti-competitive behaviour in a specific market. In some cases, this solution could be to disclose data; in others to establish Chinese wall between various activities (as in the case of Amazon marketplace on one side, Amazon e-seller on the other).

We are convinced that privacy issues should be taken into account by authorities in markets' analysis notably if such elements become a differentiated factor on which firms compete each other. Regarding this specific aspect of data, empowering consumers through the creation of specifically designed property and portability rights and imposing more transparency on data policy of economic actors, as laid down in the EU's General Data Protection Regulation (GDPR), sounds good. Indeed, such measures should reduce asymmetries of information between data producers and data users, allow to benefit from data disclosure and limit its drawbacks (notably regarding privacy threat), promoting a fairer competition.

Corresponding author: Claire Borsenberger (claire.borsenberger@laposte.fr)

References

Borsenberger C., D. Joram, O. Klargaard and P. Regnard (2016), "Personal Data and Privacy Issues and Postal Operators Stand", in M.A. Crew and T.J. Brennan (eds), *The Future of the Postal Sector in a Digital World*, Springer International Publishing, pp. 261-270.

Borsenberger C., M. Hoang and D. Joram (2019), "Open-data: a solution when data constitutes an essential facility?", to be published in P.L. Parcu, T. Brennan and V. Glass (eds), New Business and Regulatory Strategies in the Postal Sector.

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, UK: Oxford University Press.

Autorité de la concurrence and Bundeskartellamt (2016), *Competition Law and Data*.

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 volume 3, Article n° 1376.

de Montjoye Y.-A., Radaelli L., Singh V. K., Pentland A. S. (2015), "Unique in the shopping mall: On the reidentifiability of credit card metadata", *Science*, vol. 347, Issue 6221, pp. 536-539.

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.

Ezrachi and Stuckes (2017), "Artificial Intelligence & Collusion: When Computers Inhibit Competition", *University of Illinois Law Review*, Vol. 2017-5, pp. 1775-1809.

Graef, Inge (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., and M.E. Stucke (2015) "No mistake about it: The important role of antitrust in the era of big data", *The Antitrust Source*, April.

Isaac, Henri (2016), *Données, valeur et business model*, Les Cahiers Scientifiques de la Chaire IESO, 2016, n°21.

Lambrecht, Anja and Catherine E. Tucker (2015), "Can Big Data Protect a Firm from Competition", December.

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, Nils-Peter and Achim Wambach (2015), "On Big Data and Its Relevance for Market Power Assessment", *Journal of European Competition Law & Practice*.

Sokol, D., and R. Comerford (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*, Cambridge University Press, pp. 293-316.