Journal of Research in Health Sciences
ISSN: 2228-7809 | https://jrhsjournal.com

Consumer Trust and Regulatory Tensions in the European Commission's Data and AI Framework

Dr. Lucas Ribeiro[1]*, Dr. Mariana Costa[2]
[1]Department of Public Health, University of São Paulo, São Paulo, Brazil
[2]Faculty of Medicine, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Abstract

The Commission's plans to shape Europe's digital future aim at increasing the generation and availability of data. The strategy to achieve this goal is underpinned by the belief that the GDPR generates the trust required to drive more demand for data-driven technologies, as well as more data-driven competition and innovation, an outcome purportedly compatible with the protection of individuals' personal data. This article exposes the flaws in this rationale, noting that the GDPR neither generates trust nor stimulates data-driven competition. Crucially, on the most fundamental level, the protection of personal data and the promotion of consumer welfare in data-driven markets pull in opposite directions. Since the GDPR is both failing to protect individuals against the privacy risks posed by current technologies and serving as a regulatory barrier to entry that favours established dominant platforms, the EU should reconsider the current personal data protection mechanisms to arrive at normatively consistent outcomes capable of affording effective data privacy protection and improving the competitiveness of EU firms.

Keywords – GDPR – Competition – Data Strategy – Big Data – Data Privacy

1 Introduction

The European Commission (the 'Commission') recently announced its plans to shape Europe's digital future, releasing a Communication that sets out the European 'Data Strategy'(European Commission, 2020a) and a White Paper containing a number of policy options aimed at promoting the uptake of Artificial Intelligence (AI) (European Commission, 2020b). These two documents are intrinsically related, as achieving the aims of the former is a precondition for the development of the AI ecosystem contemplated in the latter. The refinement of AI depends on access to large volumes of data to train the underlying algorithms. However, in spite of the dramatic growth in data generation caused by technological progress, data is not widely accessible. As the Commission notes, "[t]he value of data lies in its use and re-use"; however, currently, "there is not enough data available for innovative re-use, including for the development of artificial intelligence" (European Commission, 2020a, p. 6). Accordingly, the Data Strategy seeks to create a single market for data, where data can be readily accessed, traded and shared (European Commission, 2020a, pp. 4–5). The Commission's goal is to increase "the use of, and demand for, data and data-enabled products and services throughout the Single Market" (European Commission, 2020a, p. 1), in a bid to enable the creation of "new products and services based on more accessible data"(European Commission, 2020a, p. 5).

The Commission's strategy is underpinned by the belief that the General Data Protection Regulation (GDPR) promotes consumer trust and thus leads to greater engagement with data-driven products and technologies, more data, and more data-driven innovation. There is also the idea that the GDPR can be used to "enable novel data flows and foster competition" (European Commission, 2020a, p. 10). Generally, the Commission sees personal data protection and the promotion of competition as two compatible goals, having full faith in the "creation of European data pools enabling Big Data analytics and machine learning, in a manner compliant with data protection legislation and competition law, allowing the emergence of data-driven ecosystems" (European Commission, 2020a, p. 5).

To be sure, the notion of a strict data protection regime being a tool capable of generating digital trust and promote competition is nothing new. This idea served as a one of the main justifications to pass the GDPR, and has been widely shared by EU institutions and officials (Albrecht, 2016; Article 29 Data Protection Working Party, 2014; European Commission, 2012, 2014)

This paper argues that the Commission's plans to shape Europe's digital future are underpinned by rhetoric instead of facts, exposing a European reality in which neither the protection of individuals against the risks posed by big data analytics nor the promotion of data-driven competition to the benefit of consumers is achieved. Section 2 explores the link between the GDPR and trust. Section 3 assesses the GDPR's ability to promote competition and the compatibility between the protection of personal data and the promotion of consumer welfare in data-driven markets.

Section 4 analyses whether the GDPR strikes an acceptable balance between the protection of individuals' data privacy and the stimulation of data-driven innovation and competitiveness. Section 5 concludes.

2 The GDPR/Digital Trust link

Consider the following narrative. Survey after survey the majority of participants say that they are concerned about how companies use their personal data. If consumers realise that their personal data is compromised, mined for constant analysis or subject to other privacy risks or violations, their trust is bound to be diminished, or perhaps lost. Without trust, some consumers are likely to refrain from engaging with data-driven technologies, opting instead for more traditional or analogue products or commerce channels. The GDPR was introduced to improve individuals' ability to control their personal data and provide them with 'efficient and operational means' to ensure they are fully informed about how their persona data is used (Reding, 2012, p. 124). With this empowerment, individuals now can see for themselves whether the data-driven products and services of their choice are consistent with their privacy preferences. As a consequence, their trust in and engagement with data-driven technologies is set to rise.

The link between a strong data protection regulatory framework such as the GDPR, the generation of trust and increased demand for data-driven products and services, as narrated above, rests on the assumption that individuals are 'digitally literate.' In fact, as a report prepared for the European Network and Information Security Agency (ENISA) observes, one of the main required inputs for trust is users' knowledge of online privacy (Castelluccia et al., 2011). Regrettably, this assumption is not premised on reality, and the GDPR contains no discussion, let alone strategy, to improve user knowledge on data privacy. Instead, it insists in the 'notice and consent' model - largely proved to be 'broken' - as a key data protection tool (Kuner et al., 2012, p. 48).

For consent to be valid, it must be informed. In online contexts, this means that individuals must carefully read the 'notice' (typically a 'privacy policy') of their service providers (e.g. a browser, a website, a smartphone app or an IoT device) to understand how their personal data is collected and used, assess potential privacy risks, and on this basis provide or deny consent to the processing of their personal data.

In reality, privacy policies discourage individuals from understanding the data protection implications of using a given service, thus impeding informed choices. Privacy policies are notorious for taking too much time to read. One study showed that a user would take 244 hours per year, or 40 minutes a day, to read all the privacy policies of the websites he visits, which is more than half of the average time users spend on the Internet (McDonald & Cranor, 2008, p. 563). The same study shows that if users actually read all such policies, this would entail USD 781 billion in opportunity costs (McDonald & Cranor, 2008, p. 564). Moreover, privacy policies are typically written in a way that requires a sophisticated level of reading comprehension that the average user does not have (Schraefel et al., 2017, p. 28). As a result, users rarely read privacy policies prior to using a service or visiting a website, and even if they did, most of them would not be able to understand them. Crucially, privacy policies stand in the way of users' primary task, which is accessing their chosen service. As Schraefel et al. explain, "[w]e click the "agree" button because clicking it gets rid of the screen so that we can get on with posting our cat video or uploading a draft of our paper to a co-editing site or synchronizing our calendar with a cloud service" (Schraefel et al., 2017, p. 28).

In this setting, it is hard to imagine how individuals can trust that their service providers afford a level of data protection that is consistent with their preferences, as they know nothing about, let alone understand, their data processing and handling practices. As a result, problematic information asymmetries become entrenched. Users remain in the dark as to the extent to which their personal data is protected. Conversely, data holders know exactly the scope and pervasiveness of their data processing operations, derive valuable insights from them, and put personal data to a number of uses in furtherance of their business interests.

3 The GDPR as a Driver of Competition and the (In)compatibility between Personal Data Protection and Consumer Welfare

Very few scenarios in which the GDPR fosters competition can be identified. A number of tech giants have consolidated their dominance largely due to their ability to collect, process and reuse vast amounts of personal data. Greater access to data allows for the conduction of more experiments and the refinement of algorithms, thus resulting in substantial product improvements (Llanos, 2018, p. 18). For example, based on search query data and other information provided by users, Google's search engine algorithms are able to render and increase the relevance of the search query results (Stucke & Grunes, 2016, pp. 172–174). Similarly, based on the data gathered from user-generated content and user interactions, Facebook's social network algorithms can increase the relevance of social network engagement, suggested friends or suggested interests that are shown to users (Llanos, 2018, p. 7). Crucially, both companies use that data to improve service personalisation and also render targeted search-based and display advertisements. Amazon, in turn, processes data

surrendered by users during their interaction with its platform (for example, browsing data, which reveal habits, interests and preferences) to tailor recommendations and deals, thereby driving more sales (Mangalindan, 2012). These platforms use the personal data they collect for multiple additional purposes which are conflated in their privacy policies, which enables them to entrench their dominance and leverage their data advantage onto related markets (Boutin & Clemens, 2017, pp. 4–5). Competitors without access to the same scale and scope of data cannot realistically challenge these incumbents, as they lack the necessary raw material to train their algorithms and thereby make their services more compelling. Dominant platforms' data advantage thus becomes an insurmountable barrier to entry. In this context, if properly enforced, the purpose limitation principle could "lead to a 'soft' break-up of dominant digital firms" (Ryan & Lynskey, 2019, p. 8). By requiring that these platforms have a separate legal basis for each data processing operation conducted in furtherance of legitimate, predictable and clearly pre-defined purposes, their ability to use personal data for disparate, incompatible purposes, in the way that cements their dominance, would be dramatically reduced. As a result, the aforementioned barrier to entry would be mitigated, thereby rendering the relevant markets more contestable. Similarly, by allowing consumers to port their data between data-driven services, the right to data portability can reduce lock-in effects and facilitate switching (Crémer et al., 2019, p. 8), especially if data mobility is enabled (Furman et al., 2019, p. 65).

The GDPR's limited ability to stimulate data-driven competition is due to the fact that, on the most fundamental level, there is an inherent tension between the protection of personal data and the promotion of data-driven innovation, an essential aspect of consumer welfare. This tension can be seen by comparing the main goals of the data protection and competition fields.

Broadly speaking, competition law seeks to protect the competitive process in the internal market.[1] The scope of this protection is far from settled. Suffice it to say here that as a consequence of the influential 'antitrust revolution' brought about by the Chicago School in the US in the 70-80s, consumer welfare has gained significant preponderance as a standard against which harms to competition are determined, thus becoming one of the– if not *the* - main goal of EU competition law. In EU competition policy, consumer welfare refers to the benefits derived from the competitive process for consumers in the form of lower prices, better quality, more choice and greater innovation (see Case C-209/10, *Post Danmark*). These benefits have a different weight depending on the market at hand. Whereas price tends to be the most significant competition parameter in commodity markets, innovation-driven considerations are typically salient in high technology markets. This is reflected, for example, in the 2010 Regulation exempting R&D agreements from the application of Art. 101(1) TFEU, where it is stated that R&D can bring benefits to consumers in the form of improved products or services or the 'quicker launch' of them. Similarly, the Commission's decisional practice features many cases underpinned by the likely negative effect of a practice on the process of innovation, that is to say, the development and introduction into the marketplace of new products, as well as the improvement of the existing ones (see *inter alia Microsoft* (tying), *Microsoft* (Internet Explorer), *Google Shopping, Google Android*).

Data-driven innovation relies on the use of information "from improved data analytics to develop improved services and goods that facilitate everyday life of individuals and organisations, including SMEs" (European Commission, 2014, p. 5). Premised on the 'quantity over quality of data' philosophy inherent to big data (van der Sloot & van Schendel, 2016, p. 120), data-driven innovation has enabled the launch into the marketplace of ground-breaking products services, such as applications that improve students' learning assessments, medical monitoring technology that improve patient outcomes, and solutions that provide data-driven intelligence and insights for small businesses (Software & Information Industry Association, 2013, pp. 13–15). In addition, data-driven innovation has been pivotal for the development and improvement of search engines, social media, ecommerce and online advertising, also enabling the 'smart grid' and efficiencies in traffic management, retail, logistics and fraud detection (Tene & Polonetsky, 2012, pp. 248–250). Aside from businesses that engage in data-driven innovation, (businesses that use 'data-driven decision-making' reportedly enjoy a 5-6% increase in productivity; see Tene & Polonetsky, 2012, p. 243) consumers are the main beneficiaries of the resulting innovative outcomes, not least when they take the form of 'zero-priced' products and services that are exchanged for personal data, as low prices are traditionally seen as a 'boon to consumers'. Therefore, if not wielded to engage in anti-competitive conduct, Big data and associated technologies, the enablers of data-driven innovation and source of multiple benefits for consumers, are in line with the promotion of consumer welfare.

Conversely, data protection, as both a field of law and policy and a fundamental right, is the product of early European discussions on the privacy-related threats posed by information communication technologies (ICT) (Bygrave, 2010). Data protection law aims to prevent harmful

[1] As Protocol 27 on the internal market and competition makes clear: 'the internal market as set out in Art 3 of the Treaty on European Union includes a system ensuring that competition is not distorted'.

consequences on individuals' fundamental rights and freedoms - such as the right to privacy, self-determination, non-discrimination, autonomy, integrity, dignity and reputation - that may ensue from the misuse of personal data (Bygrave, 2002; Purtova, 2012; Wachter & Mittelstadt, 2019). The concept of personal data has been broadly construed by the Article 29 Working Party and the Court of Justice of the European Union, in order to ensure a high level of protection of individuals' fundamental rights and freedoms. This protection is triggered in the form of oversight and control over how personal data is collected and processed. In particular, the GDPR provides for an array of principles, mechanisms and rights that seek to prevent unnecessary data collection, disclosure and transfer of personal data, and ultimately impede individuals from being unduly identified or singled out. Whilst the GDPR's success in attaining its goal of enhancing the protection of personal data is anything but unquestionable (see section 4 below), it is undeniable that many of its core principles and stringent requirements are incompatible with the core tenets of big data (Zarsky, 2016, p. 996).

Think of big data and the purpose limitation principle. Big data entails the combination and re-usage of large volumes of data collected in diverse contexts to extract hidden or unpredictable inferences and correlations for purposes which are typically unknown at the time of data collection. The purpose limitation principle, conversely, was designed to set the boundaries within which personal data collected for a particular purpose may be subsequently used, thereby inhibiting 'mission creep' which "could otherwise give rise to the usage of the available personal data beyond the purposes for which they were initially collected" (Article 29 Data Protection Working Party, 2013, p. 4). By limiting both the collection of personal data and its re-usage, the purpose limitation notion prevents the 'datafication of everything' and the threats this entails (Hildebrandt, 2015, p. 205).

Relatedly, since more data enhances the likelihood of valuable insights being found, not only is data directly useful for the purpose for which it is processed collected and retained by default, but also data the usefulness of which has expired (i.e. it is no longer necessary for such purpose), as well as data of mere potential utility (Rouvroy, 2016, p. 5). This practice sits at odds with the data minimisation and storage limitation principles, which limit data collection to what is strictly necessary in relation to the specific purpose that legitimise the processing and require that said data be deleted after fulfilment of such purpose. Additional tensions between the GDPR and big data practices can be found in the provisions governing consent, data accuracy, the protection of special categories of data, automated decision-making including profiling, data-protection by design, and more generally, the distinction between personal and non-personal data which determines the GDPR's scope of application.

It follows that a strict data protection regulatory framework such as the GDPR impedes "the flow of personal data, as well as the ways it [can] be analy[s]ed and used" (Zarsky, 2016, p. 1002). As a result, the GDPR is liable to compromise the growth of big data, the scope of data-driven innovation, and the ensuing benefits consumers can derive from it. This tension, which is entirely overlooked in the Commission's Data Strategy, manifests itself in a number of scenarios that lead to normatively inconsistent outcomes. Two of these scenarios are presented below: the imposition of a data-sharing obligation on dominant undertakings under Article 102 TFEU and data-driven efficiency defences in merger control.

Article 102 TFEU can be relied upon to restore competition by removing the bottleneck for data access by the incumbent's competitors, provided that the incumbent holds a dominant position in the relevant market and the refusal to give access to data constitutes an abuse of that dominant position. The idea is that access to the incumbent's data by competitors is likely to enable them to innovate and improve their services, compete on the merits and reduce the extent of the incumbent's data advantage. In this context, the French Competition Authority (*Authorité de la Concurrence*) ordered GDF Suez in 2014 to grant competitors access to parts of its database of clients, which would ensure that competing gas suppliers could compete more effectively with GDF by enabling them to better inform customers of alternative offers available to them (Autorité de la Concurrence, 2015). In particular, with regard to customers who had a contract with GDF Suez for the supply of gas under regulated tariffs established pursuant to GDF Suez' public service obligation, GDF was forced to provide its competitors with customer personal data including names, home address, fixed telephone numbers and consumption profiles. This measure was imposed after a competitor, Direct Energie, complained that GDF Suez's large datasets about customers on regulated tariffs gave it an unmatchable advantage for maintaining its dominant position in the gas market and acquiring new customers in the electricity market.

The French Competition Authority implemented the data-sharing obligation imposed on GDF Suez subject to an opt-out system, under which customers had to actively impede other gas suppliers from gaining access to their personal data. This is a problematic solution from a data protection standpoint. It has been proved over and over again that consumers rarely change default settings. In addition, to promote individuals' informational self-determination, it can be safely argued that consent was the most appropriate

ground to legitimise the data-sharing.[2] However, whilst it can be maintained that an opt-out mechanism satisfied the requirements of consent under the Data Protection Directive, this is not the case under the GDPR, which requires a 'clear affirmative action' signifying the data subject's agreement to the processing of his/her personal data. Moreover, in addition to reliance on a lawful ground for processing, the data quality requirements set out in Article 5(1) GDPR need be met, including the purpose limitation principle. Under the first prong of this principle (i.e. purpose specification), data subjects ought to be informed of the fact that their personal data will be shared with third parties and consequently processed for a new purpose.

In merger review, when a proposed concentration is assessed to determine whether it will significantly impede effective competition, the merging parties are afforded the possibility to claim and prove that the efficiencies stemming from the concentration outweigh any likely anti-competitive effects derived therefrom. When data-driven efficiencies arising from the combination of the merging parties' datasets that include personal data are put forward, an inevitable clash between the goals of personal data protection and the promotion of data-driven competition and innovation emerges. One of such defences was raised in *Microsoft/Yahoo! Search Business*, where Microsoft claimed that the transaction was underpinned by the fact that scale was essential to effectively compete in the search and search advertising markets. The Commission observed that "scale is an important element to be an effective competitor", and that the majority of respondents to the market investigation considered that Microsoft did not have enough traffic volume to be an attractive alternative to Google (*Microsoft/Yahoo! Search Business,* para. 153). In addition, it found that "the effects of scale [were] likely to allow the merged entity to run more tests and experiments on the algorithms in order to improve its relevance" (para. 223). The Commission ultimately approved the merger, as it predicted that the merged entity would enjoy greater scale of data and therefore would be able to improve its algorithms through trial and error, thereby exerting more competitive pressure on Google. Whilst the approval of the transaction makes sense on competition grounds, it is unsatisfactory from a data protection perspective. Users of both Microsoft's Bing and Yahoo! could not have anticipated that their personal data was going to be combined with other datasets to derive more, oftentimes sensitive inferences about them for the provision of search and search advertising services, nor did they have the opportunity to challenge this combination. Data-driven efficiency defences were also put forward in *TomTom/Teleatlas*, and although not expressly claimed by the merging parties, the attainment of efficiencies arising from the combination of datasets containing personal data was the underlying rationale of the *Facebook/WhatsApp* merger.

4 Is there an acceptable balance between the protection of individuals and the promotion of innovation?

The argument can be made that EU data protection law, as conceived in the GDPR, is the framework which more adequately balances the protection of individuals' data privacy with the promotion of data-driven innovation and competition. To corroborate the validity of this argument, it must be determined whether, and if so to what extent, the GDPR achieves its goal of protecting individuals' data privacy (4.1) and is actually conducive to more data-driven innovation and competition in Europe (4.2). As this section demonstrates, this argument is not supported by facts.

4.1 Does the GDPR adequately protect individuals?

The GDPR's mechanism to protect individuals – i.e. preventing *ex ante* 'unnecessary' data collection and processing with an aim to impede that individuals be unduly identified or singled out – is both outdated and ineffective.

Regard being had to ubiquitous connectivity, the rise of big data, online tracking, real-time bidding, cloud computing and the Internet of Things, it is hard to convincingly argue that the GDPR has prevented or is likely to prevent 'unnecessary' data collection and processing in reality, not least given that data controllers have no incentive to reduce the scope of their data processing practices. In addition, because of the GDPR's "many open and fuzzy norms, [controllers] can easily argue that what they do is 'necessary' for the purposes they define themselves with usually less than razor-sharp precision, until, in rare cases, some supervisory authority stops them"(Koops, 2014, pp. 254–255). Data protection enforcement is unlikely to improve. Empirical research has shown that data protection supervisory authorities across the EU are severely understaffed and under-resourced. According to Ryan and Toner, "[t]wo years after the GDPR was first applied, the principles of data protection remain almost entirely unenforced online", given that "European Governments are not providing technical staff and budgets for major legal

[2] Strictly speaking, compliance with a data-sharing obligation imposed in legal proceedings is consistent with the legal ground for processing contemplated in Article 6(1)(c) GDPR. However, whilst reliance on this ground is a legally acceptable solution, it can be hardly argued that bypassing consent in this way promotes data subjects' rights and informational self-determination.

contests to their national data protection authorities" (Ryan & Toner, 2020, p. 2).

Data controllers could argue that the magnitude of data collection is not necessarily a problem, as personal data can be anonymised, thereby impeding the undue identification of the individuals to which said data relate. This, however, is also unconvincing. Given the state-of-the-art in data processing technologies and the amounts of data available for analysis, achieving irreversible anonymisation is no longer possible (this is a well-documented reality; see Ohm, 2009, p. 1742; Schwartz & Solove, 2011, p. 1877; Tene & Polonetsky, 2012, p. 258). Examples of the failure of absolute anonymisation are abundant. Early in 2008, the film rating records of 500,000 Netflix subscribers were re-identified using the public Internet Movie Database (Narayanan & Shmatikov, 2008). Similarly, it was shown in 2015 that knowledge of four random pieces of information was sufficient to re-identify 90% of individuals in an anonymous dataset containing three months of credit card transactions by 1.1 million users. Tellingly, knowledge of one additional transaction increased the risk of re-identification by 20% (De Montjoye et al., 2015). More recently in 2019, researchers published a method to correctly re-identify 99.98% of individuals in anonymised datasets with just 15 demographic attributes (Rocher et al., 2019).

Furthermore, the GDPR is ill-equipped to protect individuals against the risks posed by inferential analytics. Looking for the 'what' without knowing the 'why', big data analytics yields connections and correlations that are both unexpected and previously unknown. For example, it can be known that a person who buy diapers is more likely to also buy beer (Siegel, 2013, p. 117), but it cannot be known why this is actually the case. On the basis of these correlations, inferences about individuals' and groups of individuals' behaviour, preferences and private lives are made. These inferences "can be used to nudge and manipulate us" (Wachter & Mittelstadt, 2019, p. 13), typically for financial gain. Early in 2010, then Google's CEO Eric Schmidt claimed that "[individual targeting] technology will be so good it will be very hard for people to watch or consume something that has not in some sense been tailored for them." In digital marketing, rather than adapting supply to individuals' spontaneous wishes, the goal is to adapt such wishes to what is being offered by tailoring sales strategies to each individual's interest profile, thus depleting "limited resources of will-power"(Calo, 2013, p. 1031). Giant retailer Amazon patented an 'Anticipatory Shipping' software that predicts what buyers are going to buy and ships products to their doorstep, even before placing the order ('Amazon Patents "Anticipatory" Shipping — To Start Sending Stuff Before You've Bought It', n.d.). These data analytics tools are highly damaging to individuals' self-determination and identity: "[r]ather than deciding for yourself 'who am I' and 'what do I want' […], big data creates the risk [of] turning this into being told 'who you are' and 'what you want'(Moerel, 2014, p. 9). However, data subjects' right to know about, rectify, delete, object to and port personal data[3] are considerably attenuated in respect to inferences, typically requiring a "greater balance with the controller's interests (e.g. trade secrets or intellectual property) than would otherwise be the case" (Wachter & Mittelstadt, 2019, p. 6).

4.2 Does the GDPR have a positive impact on the EU's innovation and competitiveness?

Innovation depends on a magnitude of factors, such as the degree of government intervention, start-up culture, the quality of higher education (particularly elite university research), public support for the formation of innovation clusters (such as Silicon Valley) and the choice of instruments for public funding (Forge et al., 2013, pp. 7–9). The extent to which the regulatory environment enables innovation paths is another aspect of great importance. Insofar as data protection laws place restrictions on the collection, usage and re-usage of data, the degree to which they are restrictive or permissive is likely to have an impact on data-driven innovative outcomes. An assessment of and comparison amongst the data protection regulatory frameworks of the EU, the US and China and their success in data-driven sectors lend support to this notion. According to Castro *et al.,* "[b]y imposing stringent restrictions on the collection and use of data, the GDPR makes it more challenging for businesses to use the data consumers are creating". They conclude that "the EU's regulatory environment creates the most restrictions on the collection and use of data, followed by the United States and China"(Castro et al., 2019, p. 42).

When innovation in data-driven sectors is compared between the EU and the US, it is clear that "not only is Europe failing to establish leadership in the internal market, it is unable to produce a presence"(Zarsky, 2015, p. 155). Whereas the US has seen the birth, growth and dominance of highly innovative data-driven platforms like Google, Amazon, Facebook, Apple (GAFA), Twitter, eBay and Uber, no EU-based firm has been remotely close to challenge them or match their scale. Moreover, under its - until recently non-existent and currently permissive - data protection regime, a number of Chinese tech companies have risen, grown exponentially and even established international presence. Baidu, Alibaba and Tencent (commonly referred to as 'BAT') are the most famous

[3] Articles 13-15, 16, 17, 20 and 21 GDPR.

examples, the last two being featured in the top 10 most valuable companies in the world in 2020 (*Most Valuable Companies in the World - 2020*, n.d.). Each of these firms has created complex ecosystems composed of multiple platforms and components (OECD, 2019, pp. 122–123, 91–95, 181–184), aided by their ability to collect and process unprecedented volumes of data (OECD, 2019, p. 24).

Correlation does not equate to causation. The US' and China's ability to succeed in data-driven innovation and give birth to tech giants is likely to be due to additional factors, such as the US's 'risk-taking culture'(Thierer, 2014) and the Chinese government's protectionism which insulated Chinese firms from competition from American platforms (OECD, 2019, p. 41). However, the fact that data-driven big players emerge in jurisdictions with lax, business-friendly and consumer-oriented data protection laws, and not in the EU, "the global gold standard in the protection of personal data" (European Commission, 2015), does suggest that the EU strict data protection regulatory framework, contrary to the Commission's rhetoric, has done and does very little to promote data-driven innovation and the competitiveness of EU firms.

Crucially, whereas the GDPR was meant to "substantially reduce the administrative burden" on controllers and processors (Reding, 2011), in reality, this has not been the case. The GDPR's strict requirements on consent, information disclosure, transparency and accountability, to name a few, involve substantial record-keeping and red tape. Also, research has shown that the GDPR requires companies to build a dedicated data management capability which involves the re-designing of data-processing systems (Jakobi et al., 2020, p. 265). All of the foregoing involves costs which not every company is able to bear. It has been estimated that an average firm of 500 employees must spend around USD 3 million to comply with the GDPR (International Association of Privacy Professionals, 2019). Yet, SMEs are expected to fulfil their obligations and "manage their data flows and data processes to the same extent as bigger and better resourced organisations" (ENISA, 2016, p. 16). This reality tilts the playing field in favour of incumbents, and the GDPR effectively becomes a barrier to market entry. As Facebook Chief Operating Officer Sheryl Sandberg observed, "it's actually easier for big companies like Facebook, or other big competitors, to put in place things that adhere to regulation than it is for startups. If I think back to Facebook 10 years ago, GDPR would have been much harder for us then than it was now"(Schechner, 2019). As a result, instead of enabling EU startups and SMEs to compete, the GDPR has reinforced the dominance of American platforms like Google, Amazon and Facebook (Schechner, 2019), and effect which was predicted before its entry into force (Wakabayashi & Satariano, 2018).

5 Conclusions

The GDPR does not generate trust. Trust requires an understanding of what is at stake, that is, the level of data protection afforded by online and data-driven products and services. But data subjects cannot realistically be expected to read and understand every single privacy policy they are confronted with. We have clicked, continue to click, and will remain clicking 'accept' without reading and understanding absolutely anything. The GDPR does not contribute to remove information imbalances concerning data privacy. Moreover, the GDPR's ability to promote data-driven competition is highly limited, not least given that the protection of personal data, as currently contemplated, and the promotion of data-driven innovation pull in opposite directions.

Contrary to the rhetoric exposed in this paper, it is apparent that the GDPR stands in the way of the Commission's goal. Nor is the GDPR succeeding in protecting individuals against the privacy and associated risks posed by current data processing technologies. Therefore, it is submitted that the time has come for the EU institutions to engage in honest debate on and reconsideration of the current personal data protection mechanisms, with a view to afford actual, effective protection of individuals and arrive at consistent normative outcomes across the fields of data protection and competition that are capable of improving the competitiveness of EU firms.

Acknowledgements

This research was conducted within the context of the Privacy-Aware Cloud Ecosystems (PACE) project funded by the Engineering and Physical Sciences Research Council (EPRSC), reference: EP/R033439/1.

References

Albrecht, J. P. (2016). How the GDPR will change the world. *Eur. Data Prot. L. Rev.*, *2*, 287.

Amazon Patents "Anticipatory" Shipping—To Start Sending Stuff Before You've Bought It. (n.d.). *TechCrunch*. Retrieved 24 July 2020, from https://social.techcrunch.com/2014/01/18/amazon-pre-ships/

Article 29 Data Protection Working Party. (2013). *Opinion 03/2013 on Purpose Limitation* (00569/13/EN WP 203).

Article 29 Data Protection Working Party. (2014). *Statement on Statement of the WP29 on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU*. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp221_en.pdf

Autorité de la Concurrence. (2015, February 18). *Autorité de la concurrence issues Urgent Interim Measures ordering GDF, Incumbent Gas Supplier, to Grant its Competitors Access to its Client Database—ECN Brief—European Commission*. ECN Brief. /multisite/ecn-brief/en/content/autorit%C3%A9-de-la-concurrence-issues-urgent-interim-measures-ordering-gdf-incumbent-gas

Boutin, X., & Clemens, G. (2017). Defining 'Big Data' in Antitrust. *Competition Policy International: Antitrust Chronicle*, *1*(2), 22–28.

Bygrave, L. A. (2002). *Data Protection Law: Approaching its Rationale, Logic and Limit*. Kluwer Law International, The Hague/London/New York.

Bygrave, L. A. (2010). Privacy and data protection in an international perspective. *Scandinavian Studies in Law*, *56*(8), 165–200.
Calo, R. (2013). Digital market manipulation. *Geo. Wash. L. Rev.*, *82*, 995.

Castelluccia, C., Druschel, P., Hübner, S., Pasic, A., Preneel, B., & Tschofenig, H. (2011).

Privacy, accountability and Trust-Challenges and opportunities. *ENISA.[Online]. Available: Http://Www. Enisa. Europa. Eu/Activities/Identity-and-Trust/Library/Deliverables/Pat-Study/at Download/FullReport*.

Castro, D., McLaughlin, M., & Chivot, E. (2019). *Who Is Winning the AI Race: China, the EU or the United States?* https://www.datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/

Crémer, J., de Montjoye, Y.-A., & Schweitzer, H. (2019). *Competition Policy for the Digital Era*.

De Montjoye, Y.-A., Radaelli, L., & Singh, V. K. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, *347*(6221), 536–539.

ENISA. (2016). *Guidelines for SMEs on the security of personal data processing*. file:///Users/tomas/Downloads/WP2016%203-2%206%20Data%20Controllers%20Risk.pdf

European Commission. (2012). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—Safeguarding Privacy in a Connected World. A European Data Protection Framework for the 21st Century* (COM(2012) 9 final).

European Commission. (2014). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—Towards a thriving data-driven economy* (COM(2014) 442 final). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0442&from=GA
European Commission. (2015, January 28).

European Commission Statement, Vice-President Ansip and Commissioner Jourová: Concluding the EU Data Protection Reform Is Essential for the Digital Single Market'. European Commission. https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_15_3801

European Commission. (2020a). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—A European Strategy for Data* (COM(2020) 66 final).

European Commission. (2020b). White Paper on Artificial Intelligence—A European Approach to Excellence and Trust. *COM(2020) 65 Final*.

Forge, S., Blackman, C., Goldberg, I., & Bagi, F. (2013). *JCR Technical Reports—Comparing Innovation Performance in the EU and the USA: Lessons from Three ICT Sub-Sectors* (Report EUR 25961 EN).

Furman, J., Coyle, D., Fletcher, A., McAuley, D., & Marsden, P. (2019). *Unlocking Digital Competition. Report of the Digital Competition Expert Panel*.

Hildebrandt, M. (2015). *Smart technologies and the end (s) of law: Novel entanglements of law and technology*. Edward Elgar Publishing.

Independent German Federal and State Data Protection Supervisory Authorities. (2019). *Report on Experience Gained in the Implementation of the GDPR*. https://www.datenschutzkonferenz-online.de/media/dskb/20191213_evaluation_report_german_dpa_s_clean.pdf

International Association of Privacy Professionals. (2019). *IAPP-EY Annual Governance Report 2018*. https://iapp.org/resources/article/iapp-ey-annual-governance-report-2018/

Jakobi, T., von Grafenstein, M., Legner, C., Labadie, C., Mertens, P., Öksüz, A., & Stevens, G. (2020). The Role of IS in the Conflicting Interests Regarding GDPR. *Business & Information Systems Engineering*, 1–12.

Koops, B.-J. (2014). The trouble with European data protection law. *International Data Privacy Law*, *4*(4), 250–261.

Kuner, C., Cate, F. H., Millard, C., & Svantesson, D. J. B. (2012). The challenge of 'big data' for data protection. *International Data Privacy Law*, *2*(2).

Llanos, J. T. (2018). The Data Paradox in Competition Enforcement. *TLI Think! Paper 10/2019*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3373553

Mangalindan, J. P. (2012). *Amazon's recommendation secret*. Fortune. http://fortune.com/2012/07/30/amazons-recommendation-secret/

McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy*, *4*(3), 543–568.

Moerel, E. M. L. (2014). *Big data protection: How to make the draft EU regulation on data protection future proof*. http://www.debrauw.com/wp-content/uploads/NEWS%20-%20PUBLICATIONS/Moerel_oratie.pdf?utm_source=Mondaq&utm_medium=syndication&utm_campaign=inter-article-link

Most Valuable Companies in the World—2020. (n.d.). FXSSI - Forex Sentiment Board. Retrieved 25 July 2020, from https://fxssi.com/top-10-most-valuable-companies-in-the-world

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. *In 2008 IEEE Symposium on Security and Privacy (SP 2008)*, 111–125. http://ieeexplore.ieee.org/abstract/document/4531148/

OECD. (2019). *An Introduction to Online Platforms and Their Role in the Digital Transformation*. OECD Publishing, Paris.
Ohm, P. (2009).

Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA l. Rev.*, *57*, 1701.

Purtova, N. (2012). *Property rights in personal data: A European perspective*. Alphen aan den Rijn: Kluwer Law International.

Reding, V. (2011, June 20). *Assuring Data Protection in the Age of the Internet' (SPEECH/11/452, at BBA (British Bankers' Association) Data Pro- tection and Privacy Conference, London*. European Commission - European Commission. https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_11_452

Reding, V. (2012). The European data protection framework for the twenty-first century. *International Data Privacy Law*, *2*(3).

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications*, *10*(1), 3069. https://doi.org/10.1038/s41467-019-10933-3

Rouvroy, A. (2016). 'Of data and men.' Fundamental rights and freedoms in a world of big data. *Report for the Council of Europe's Bureau of the Consultative Committee of the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data*.

Ryan, J., & Lynskey, O. (2019). *Contribution to the CMA's online platforms and digital advertising market study*.

Ryan, J., & Toner, A. (2020). *Europe's governments are failing the GDPR - Brave's 2020 report on the enforcement capacity of data protection authorities*.

Schechner, N. K. and S. (2019, June 17). GDPR Has Been a Boon for Google and Facebook. *Wall Street Journal*. https://www.wsj.com/articles/gdpr-has-been-a-boon-for-google-and-facebook-11560789219

Schraefel, M. C., Gomer, R., Alan, A., Gerding, E., & Maple, C. (2017). The internet of things: Interaction challenges to meaningful consent at scale. *Interactions*, *24*(6), 26–33.

Schwartz, P. M., & Solove, D. J. (2011). The PII problem: Privacy and a new concept of personally identifiable information. *NYUL Rev.*, *86*, 1814.

Siegel, E. (2013). *Predictive analytics: The power to predict who will click, buy, lie, or die*. John Wiley & Sons.

Software & Information Industry Association. (2013). Data-Driven Innovation, A Guide for Policymakers: Understanding and Enabling the Economic and Social Value of Data. *SIIA White Paper Series*.
Stucke, M., & Grunes, A. (2016). *Big Data and Competition Policy*. Oxford University Press.

Tene, O., & Polonetsky, J. (2012). Big data for all: Privacy and user control in the age of analytics. *Nw. J. Tech. & Intell. Prop.*, *11*, xxvii.

Thierer, A. (2014). *Embracing a Culture of Permissionless Innovation*. Cato Institute. https://www.cato.org/publications/cato-online-forum/embracing-culture-permissionless-innovation

van der Sloot, B., & van Schendel, S. (2016). Ten questions for future regulation of big data: A comparative and empirical legal study. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *7*, 110.

Wachter, S., & Mittelstadt, B. (2019). A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Colum. Bus. L. Rev.*, 494.

Wakabayashi, D., & Satariano, A. (2018, April 23). How Facebook and Google Could Benefit From the G.D.P.R., Europe's New Privacy Law. *The New York Times*. https://www.nytimes.com/2018/04/23/technology/privacy-regulation-facebook-google.html

Zarsky, T. Z. (2015). The privacy-innovation conundrum. *Lewis & Clark L. Rev.*, *19*, 115.

Zarsky, T. Z. (2016). Incompatible: The GDPR in the age of big data. *Seton Hall L. Rev.*, *47*, 995.