Big Data between privacy and competition: dominance by exploitation? Which remedies?

Fabiana Di Porto* and Gustavo Ghidini**

Introduction

The data-driven economy value-chain is based, amongst others, on a variety of data that the big platforms (1) collect through digital connections: they may be either personal or non-personal data (2). Just as bulk oil, the latter need to be refined to become useful: data must be combined and analyzed to turn into meaningful information,

---

* Fabiana Di Porto is Associate Professor of Law at University of Salento, Lecce (Italy), co-Director of the Journal “Concorrenza e mercato. Competition, Regulation, Consumer Welfare, IP” and Member of the Board of the Academic Society for Competition Law – ASCOLA.
** Gustavo Ghidini is Professor of Law at LUISS University of Rome and University of Milan (Italy), Director of the Observatory on IP, Competition and Communications at LUISS University, and co-Director of the Journal “Concorrenza e mercato. Competition, Regulation, Consumer Welfare, IP”.

1 By online platforms we mean those “key enablers of digital trade”, inclusive of both stricto sensu platforms (such as e-commerce marketplaces, software application stores and social media) and search engines, that are qualified as “quasi gatekeepers to markets and consumers” by the proposed EU Regulation, COM(2018) 238 fin. on “Promoting fairness and transparency for business users of online intermediation services”, of 26.4.2018, at p. 1 (available at: http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51803).
knowledge, allow action and finally implementation (it is the so-called big data pyramid, a concept elaborated from N. Henry, 1974 (3)).

Particularly in Europe, also at the level of EU institutions, the awareness has matured that "market power" can be, and is gained by collecting and processing huge amounts of personal (and non-personal) data, including sensitive data, of users acquired through the web (4).

In particular, as regards information society services, market power is gained by collecting and aggregating user data to the extent that the same data are utilized as the basis for providing additional services. Data is collected by offering online services, whether at a charge or (apparently: see below) free of charge (5), through which user data of varied nature are collected to form: e.g. personal, identification and travel data or information concerning interests and preferences.

Market power is thus generated as a large volume of these data, after they have been collected and processed (see below on profiling), accrue to the big platforms, making it difficult, if not impossible for any competitor wishing to enter the market to offer alternative or, in any case, really competitive services. For instance, in the market for general online search services, Google’s dominant position has been established(6), amongst others, due to “the existence of barriers to expansion and entry”, which impinges on the volume of individual queries transmitted on the “free” side of the platform; but also on the

---

4 According to M. Porter (2001) Strategy and the Internet, Harvard Business Review, March 2001: “Internet technology provides better opportunities for companies to establish distinctive strategic positionings than did previous generations of information technology.” (also positing that data are a tool to foresee and influence consumers' behaviours on the markets, thus becoming a product offered on the market).
lack of countervailing buyer power, which is due to “the infrequency of user multi-homing” (7). The conclusion holds even though “general search services are offered free of charge”, precisely because users “contribute to the monetization of the service by providing data with each query”.(8)

In this framework, we wish to focus on the “re-combination” of data (9) (‘processing’ at large) to imply those activities, usually done by algorithms and thanks to huge computational capacities, through which data are analyzed, trained, and more broadly used to perform a given task (e.g. develop a product or service, be it voice recognition, web searching, and let the many applications provided through platforms operate). Re-combination implies that data belonging or pertaining to an individual’s preferences are collected to shape a behavioral profile or perhaps several profiles of each individual — multiple personalities?!...— be they cultural, political, sexual, and more broadly consumerist.

These profiles are thus created by combining all available data that pertain to her and, most importantly, are commercially exploited for selling ‘targeted’ advertising, products, services and/or for reselling to other companies. Indeed, not only are profiles created through the re-combination of personal and non-personal data, but they are at times resold to other companies taking advantage of loose contractual terms and disclosures and/or by contravening privacy rules. For instance, data on physical activity can well be merged with those on geo-localization and restaurants’ “check-ins” to create profiles that might be of

---

7 According to the Commission “only a minority of users in the EEA that use Google’s general search service as their main general search service use other general search services (a behaviour known as “multi-homing”)” (pt 306 of the Google Decision, cited above.)
8 See pt 320.
9 Or “data fusion” in the wording of A.P. Grunes and M.E. Stucke (2015), “No Mistake About It: The Important Role of Antitrust in the Era of Big Data”, 14 Antitrust Source, p. 1, at 12, meaning “linking data of diverse types from disparate sources in support of unified search, query, and analysis”, that “may yield potential uses that the consumer never envisioned”.
commercial interest for (and therefore sold to) fitness centers that seek to advertise the opening of new premises. Moreover, by using technologies such as APIs (or application programming interfaces) platforms may embed their services into third parties websites, thus collecting more off-platform data and magnifying their “identity-based network effects” (see below).

In addition, unfair trading practices pertaining to data, such as the delisting of products and services, unexplained suspension of accounts, unclear rankings, or non transparent conditions for access to and use of the data collected, are at the core of normative initiatives by the EU Commission, seeking to regulate contractual relationships between platforms and its business users (SMEs especially). (10) Finally, as recent chronicles are allegedly uncovering, deep data-sharing agreements between platforms and device manufacturers might have been used as a strategy to expand the reach and market position of a big platform, without the latter acquiring explicit consent by its users and their connections. (11)

Even when the collection refers to non-personal data, well-trained algorithms may infer from personal information. As clarified by the European Parliament, in its Resolution of 14 March 2017, thanks to algorithms of big data analytics, “sensitive information about persons can be inferred from non-sensitive data, which blurs the line between sensitive and non-sensitive data” (12). The same holds true with regards

11 See G.J.X. Dance, N. Confessore and M. LaForgia, “Facebook Gave Device Makers Deep Access to Data on Users and Friends”, The New York Times, 3.6.2018, reporting that since 2007, “Facebook has reached data-sharing partnerships with at least 60 device makers — including Apple, Amazon, BlackBerry, Microsoft and Samsung — over the last decade, starting before Facebook apps were widely available on smartphones, company officials said. The deals allowed Facebook to expand its reach and let device makers” retrieve data of both Facebook users (because they use popular features of the social network, such as messaging, “like” buttons and address books) and of their friends, including those “who have denied Facebook permission to share information with any third parties.”
12 EU Parliament, Resolution “Fundamental rights implications of big data”,
to the distinction between personal and non-personal data, provided that non-personal data fragments, once recombined, increase the possibility of re-identification and therefore of profiling (13).

Personalized digital “profiling” lays thus at the heart of the data-driven economy value-chain: while it allows digital firms (especially platforms) to capitalize on economies of scale and scope, realize efficiencies and develop tailored products and services (e.g. advertising) that may (or may seem to) maximize users’ utility, at the same time, it works as a leverage for platforms’ market power, by allowing its accumulation and magnification.

It has been recognized that large platforms by “converging control of content, access, and online distribution channels”, may “access to an immense volume of users’ personal online data” (14). Indeed, the fact that “such intermediaries currently serve as major gateways to the digital world, enables them to accumulate more data”. And given the range of services they offer, these intermediaries can “create better user profiles” (15), thus reinforcing their positions in the market (16).

The paper aims at elaborating on the competitive and consumer welfare concerns deriving from the accumulation of data, positing that market power is acquired, and dominance attained or reinforced, through “exploitation”. The latter, in particular, is understood as threefold: firstly, from the supply-side, recent European case law involving big platforms, as well as recent normative initiatives at EU

15 Id., p. 30.
level, confirm the strict links existing between re-combination of vast amounts of data and competition concerns (para. 2); secondly, from the demand-side, exploitation may also take the form of profiting from widely diffused cognitive biases affecting online consumers’ behaviors (para. 3); thirdly, the possibility to commercially exploit and trade data and digital profiles outside individual and SMEs’ control or awareness is also well-documented, raising the question of possible exploitation of existing data-protection and other laws, leading to an increase of market power or its abuse (para 4). Para. 5 discusses possible remedies to dominance-by-exploitation (and to eventual abusive practices).

2. The Supply-Side. Factors Reinforcing Market Power

Platforms are important gateways to online markets, and their business model is, as known, well captured by the two (or multiple) sided market image. The latter being characterized by data-driven network effects, its value increases exceptionally rapidly "with the number of additional users on either side, while the cost increase to provide services to additional users on either side grows increasingly slowly.” (17). With decreasing marginal costs (e.g. support functions are automated when consumers scale to millions), the more users are attracted and retained to both sides of the platforms, the greater its value. It follows that the greater the “access to high quality, variety and volumes of data”, and therefore “insight into users’ profiles/preferences”, the more “increased returns to scale, scope and network effects” will be, and therefore the so-called “data-driven competitive advantage”.

While it is acknowledged that the value of a single data is per se scarce, what allows data to become valuable is the possibility of its “reuse”. The big data value chain thus requires data to be first collected and merged with other data. That happens through the use of a range of

---

different technologies, such as tracking cookies, digital fingerprints (18) or history-sniffing techniques. However, most of the data come from activities that individuals exert using their mobile devices and through real-time sensors of the Internet of Things. (19) Once collected, data are analyzed to become meaningful information: the most widely diffused paradigm to organize big data is that of data lakes, which allows to store and quickly share different kinds of data (structured, semi-structured and de-structured). Once stored in data lakes, big data can be accessed to extract meaningful information, insight (knowledge), and thus value. Any organization hence store its data and information (or files) in its own way, depending on its business models and decisional processes.

Due to the above-mentioned network externalities and the technical characteristics of the data value chain just described, the big data ecosystem tends naturally towards market concentration, where big platforms emerge that mediate interactions among the two (or more) sides of the market. (20)

Moreover, somehow elaborating on a widely settled view, deeper and more recent analysis reveals that barriers to entry, development and access to big data exist that affect all of the steps of the value chain, starting from the very collection phase. Indeed, it is already known that strong interrelations between the initial collection phase and the following ones exist that provide considerable competitive advantages to those firms with greater capabilities of collecting digital data and/or those that can aggregate heterogeneous

---

19 See Italian Communications Authority (ICA), Big data. Interim report, dec. n. 217/17/Cons, of 8.6.2018, p. 17.
20 ICA, cited above, nt 18, p. 22-23, noting that “the presence of supply-side scale economies are a consequence of the increase of data volume and their capacity or creating increasing returns to scale, as well as economies of scope (or variety), depending, in turn, on the increasing possibility to combine a greater variety of data. That structure of costs (showing decreasing medium costs and low to null marginal costs), connected to the firm size distribution, is made even more asymmetric or skewed by the huge fixed and sunk costs for R&S activities.”
datasets in a more efficient way, and/or with better data analytics tools. Barriers to entry and expansion (be they of technological, legal and/or strategic nature) may thus operate simultaneously, reinforcing one another and making it difficult for competing firms to enter or expand in specific markets, while having the potential to create durable market power or to serve as a basis for anticompetitive conduct. (21)

More evidence is now gathered on how such barriers to entry and expansion crucially depend on the collection phase, due to the spillover effects that the creation of a barrier at the initial stage generates on the following and connected, dependent ones. (22) Thus, data fusion from different sources are observable (23) that raise concerns for competition. In particular, both in the search and social network markets one can see very similar dynamics, showing how the acquisition of the critical mass of users needed for the full exploitation of network effects and the affirmation as a platform (with stable market shares: see Figg. 1 and 2) occurred within relatively few years (from 2001-2006 for Google and 2006-2010 for Facebook). (24)

22 ICA, cited above nt 18, p. 23.
23 See The US President’s Council of Advisors on Science and Technology – PCAST (2014), Report on Big data and Privacy: A Technological Perspective, p. 21: "data fusion occurs when data from different sources are brought into contact, and new, often unexpected, phenomena emerge… Individually, each data source may have been designed for a specific, limited purpose. But when multiple sources are processed by techniques of modern statistical data mining, pattern recognition, and the combining of records from diverse sources by virtue of common identifying data, new meanings can be found." Strictly related is the "over-collection" phenomenon, also aimed to reach a critical mass of users data, which occurs when “an engineering design intentionally, and sometimes clandestinely, collects information unrelated to its stated purpose”.
24 Ibid., p. 27 and 28.
Fig. 1 – Evolution of global market shares in the online search market (%) (2001-2018)

Source: Italian Communication Authority’s elaboration on data by SEW/WebSideStory, NetApplications, NetMarketShare and StatCounter
The recent case-law by the EU Commission and the German Bundeskartellamt – to quote the most renown – confirms the strict link existing between the collection and re-combination of vast amounts of data on the one side and the risks for competition on the other.

In its Preliminary assessment to the Facebook proceeding\textsuperscript{25}, the German Federal Competition Authority (also referred to as FCA) acknowledges that by making users “\textit{choose between accepting «the

whole Facebook package», including an extensive disclosure of personal data, or not using [the app] at all”, the tech firm is using APIs technology to merge “user data obtained from third party sources” (so-called off-Facebook data26) “with data from the Facebook user’s account, even if the user has blocked web tracking in his browser or device settings”.

Off-Facebook data are those coming from third parties (eventually – although not necessarily - owned by Facebook, such as WhatsApp and Instagram) with which the platform has stipulated contracts for the use of its services (such as the "like" button or a "Facebook login” option or analytical services like “Facebook Analytics”). (27) If a third-party website has embedded such Facebook products, data are transmitted to Facebook via Application Programming Interfaces (or APIs) “the moment the user calls up that third party’s website for the first time.” (28)

In the Bundeskartellamt’s view, by doing so, due to its (alleged) dominant position in the social network relevant market, Facebook may “optimise its offer and tie more users to its network”. The FCO explains that the re-combination or “merging of the data” would increase “the «identity-based network effects» (29) .. to the detriment of other providers of social networks”. In other words, the possibility given by technologies such as APIs and the like to merging data, because of

26 See Bundeskartellamt, cited above, p. 2: «The current proceeding examines the terms and conditions Facebook is enforcing with regard to data from third party sources. These are on the one hand data generated by the use of services owned by Facebook, such as WhatsApp or Instagram, and on the other data generated by the use of third-party websites and apps. If a third-party website has embedded Facebook products such as the 'like' button or a 'Facebook login' option or analytical services such as 'Facebook Analytics', data will be transmitted to Facebook via APIs [or Application Programming Interfaces] the moment the user calls up that third party's website for the first time.»
27 See Bundeskartellamt, cited above, p. 2.
28 Ibid.
current identity network effects (based on profiling), would further increase barriers to entry and strengthen the incumbent’s position.

Concerning the effects on the advertising side of the platform, because of the multiple “user profiles generated, Facebook [would be] able to improve its targeted advertising activities”, thus “becoming more and more indispensable for advertising customers”. Even if that indispensability is not understood as essentiality in an antitrust sense, according to the FCA, it is nonetheless turning the platform into “a dominant supplier of advertising space[s].”

The Commission makes partially similar considerations in the Google search (shopping) case30. Although it does not speak explicitly of data re-combination or merger, the Commission recognizes the contribution each user supplies to “the monetization of [Google’s search] service by providing [their] data with each query” (§§158 and 32031). By incoming a query, each user “enters into a contractual relationship32”, which allows her counterpart to store and re-use her data.

That, in turn, is of huge value to the search engine, as it allows to increase its efficiency (i.e. "improve the relevance of the search service” and to “show more relevant advertising”), while at the same time strengthening barriers to entry in the advertising side of the (two-sided) platform market (§§292 and 293). In other words, receiving huge quantities of entries is what allows search algorithms to better perform their tasks (i.e. “refine the relevance of search results pages” §287); however, at the same time, as the Commission acknowledges, because the “volume” of such queries is especially essential, when it reaches a given (massive) threshold (in terms of numbers of queries), it turns into

31 “While users do not pay a monetary consideration for the use of general search services, they contribute to the monetisation of the service by providing data with each query.” (pt 320).
32 “For instance, Google’s Terms of Service provide: «By using our Services, you agree that Google can use such data in accordance with our privacy policies»” Very similar conditions apply for other search engines (pt. 158).
a barrier to entry and expansion, so that a competing general search service would need to receive a comparable volume of queries in order to compete viably. (§289)

Thus, the re-combination of personal and non-personal data (although may create economies of scale and scope) works as a leverage for digital firms’ market power (33) and, over a certain threshold of accumulation, it operates as a self-reinforcing factor transforming them into dominant platforms.

3. On the consumers’ side

On the users’ side, the Bundeskartellamt hypothesizes a loss of control “as to which data from which sources are being merged to develop a detailed profile of them and their online activities”. That, in turn, would depend on Facebook's market power, which would leave users with "no option to avoid the merging of their data". That practice is understood as Facebook’s exploitative uses of business terms, which would also constitute “a violation of the users' constitutionally protected right to informational self-determination” (p. 4).

Besides data fusion described above, data-sharing partnerships between Facebook and third parties, (allegedly) including leading device manufacturers such as Apple, Amazon, BlackBerry, Microsoft, Samsung and other Chinese companies (34), would have been signed since 2007 to help the platform “expand its reach” before its “apps were widely available on smartphones”. (35) Eventually, such partnerships would have allowed transmitting users data from the platform to third

33 Something close to what the German Federal Competition Authority (or FCA) and the French Autorité de la Concurrence had already established back in 2016: “the collection of data may result in entry barriers when new entrants are unable either to collect the data or to buy access to the same kind of data, in terms of volume and/or variety, as established companies”: Autorite de La Concurrence – Bundeskartellamt, *Competition Law and Data*, 10.5.2016 (p. 11).


parties and also permit the platform to gain control over how both consumers and third commercial digital parties alike use the data they have collected. Matching that information with Fig. 2 above would eventually help to explain the steep increase in Facebook’s market shares registered between early 2009-early 2010.

From the papers, it would also result that third parties, the manufacturers included, were allowed to “retrieve detailed information on both device users and all of their friends — including religious and political leanings, work and education history and relationship status.” (36) It would seem that no clear consent was obtained by the consumers (and their connected friends) and that data was transferred in breach of some rules or, eventually, consent decree (an issue that will be dealt with later on). (37)

Non-transparent collection of data is not directly referred to in the Google Search (Shopping) case. Here, the main demand-side factor taken into account to determine Google’s dominant position is the lack of countervailing buyer power. The latter is due to “the infrequency of user multi-homing”, which refers to the fact that “only a minority of users in the EEA that use Google’s general search service as their main general search service use other general search services.” (§ 306) The interesting thing is that the Commission states that the conclusion holds notwithstanding the fact that “general search services are offered free of charge”.(§ 320(38)

36 LaForgia and Dance, Facebook Gave Chinese, cited nt 27.
37 A point that seems to emerge from the chronicles is that by entering in deep data sharing agreements with device manufacturers, Facebook might have violated a 2011 consent decree with the Federal Trade Commission. In particular, Facebook would have “allowed the device companies access to the data of users’ friends without their explicit consent, even after declaring that it would no longer share such information with outsiders. Some device makers could retrieve personal information even from users’ friends who believed they had barred any sharing.” Ibid. (emphasis added)
38 § 320: “While users do not pay a monetary consideration for the use of general search services, they contribute to the monetization of the service by providing data with each query.”
Besides the single antitrust case, it is known that Google also makes use of API technologies (think e.g. of the Google+ button) and the many services it offers to let third companies utilize its infrastructures, while at the same time collecting more data from the digital consumer.

In general, it is true that no matter whether data are personal or non-personal - as explained above - or even anonymized, big data analytics allows for re-identification and therefore profiling. (39) (40) Moreover, the collection of data that might seem meaningless today may become valuable tomorrow thanks to re-combination: that is the reason why concepts such as data deletion or non-retention are hardly applicable in a big data environment. (41) These characteristics explain why the data collecting phase is so crucial and conditions all of the other connected stages.

One may ask then why are users so prone to release their data in such an abundant way. Research reveals that in two-sided markets, deep informational symmetries, as well as incompleteness of contracts, dominate transactions between consumers and platforms. The most diffused model for collecting data online is “notice and consent”, under which users must indicate an affirmative consent to download and use an app or service. That puts the responsibility of data privacy protection mainly on consumers on the assumption that all use and possible re-uses of data are fully acknowledged.

39 US PCAST (2014), cited nt 24, p. 38-39: “it is increasingly easy to defeat anonymization by the very techniques that are being developed for many legitimate applications of big data ... as the size and diversity of available data grows, the likelihood of being able to re-identify individuals (that is, re-associate their records with their names) grows substantially.”
40 Thus, the above-mentioned proposed Regulation on the free flow of data (above, nt 2), seeking the broadest liberalization of massive accumulation and processing of non-personal data by EU and foreign firms, "would find only limited constraints in the GDPR, and therefore contribute to magnify market power further.”: Di Porto (2018), In Praise of an Empowerment, cited at nt 13.
41 Ibid.
However, as said, data are given in exchange of no consideration and therefore do not reflect a meaningful value (i.e. price), but instead, are provided outside proper (and therefore incomplete) contracts. \(^{(42)}\) That is because when deciding to give her consent for data provision, the user has only partial knowledge about the risks and benefits deriving from her choice. Given structural informational asymmetry, not only do consumers lack the meaningful information they would need to make informed commercial decisions (that is instead available to their counterparts), but most of their behaviors, in order to be efficient, would require an extent of technical knowledge that goes far beyond the competences diffused among the population \(^{(43)}\).

In connection with this, a further powerful albeit indirect factor of market power needs to be considered. We refer to the limited awareness (disclosure profiles) and understanding (cognitive profiles) consumers tend to have of both the data they transmit to platforms (and the various legal/social implications), and their rights of access to control the use of their data, even by processing. This lack of awareness is often attributed to the users’ negligence, a modern version of the ancient pro-business mantra ‘caveat emptor’. However, the privacy statements as well as contractual terms and conditions, which signature is mandatory for the download of any apps and services, are not easy to understand, both because they are excessively lengthy, and because they are prepared (not so coincidentally) using a hardly comprehensible legal terminology.

\(^{42}\) ICA, cited above nt 18, p. 22 and 41.
\(^{43}\)Id., p. 4 and 40, explaining that (substantial) information asymmetries exist that prevent individuals to duly assess the costs of data collection: they do not dispose of information of the same quantity and quality as data collectors; there is high incertitude as per the type of collection activity and future usage of data collected; intensity of information asymmetries is high because the choice is taken in a short time frame (context dependency).
It has been estimated that a user would require 244 hours per year, i.e. more than half of the time she usually spends on the Internet, in order to read the privacy statements of all the websites she visits (44).

Moreover, as known and confirmed by recent research, most smartphone applications, increasingly present in the consumers’ daily lives, do not even require the user’s approval in respect to any personal data protection disclosures (45).

In all these typical cases, the users do not understand the purpose for which their data are being processed, granting the firms controlling their data the freedom to process them and to use them for purposes other than those for which they were acquired from the party concerned. Similarly, because the user is not familiar with the terms of processing, she is unable to exercise, or to have the relevant authorities exercise the safeguards that are theoretically in place.

Equally—and possibly even more diffusely—individual users /‘simple citizens’ so to say, are not generally aware of their rights of access and control, of the ways their data are merged to form profiles and commercially used, and of the modes of enforcing said rights. How many users, even in Europe, know of the chances opened by the new GDPR? We think in particular to the “right of access” (Article 1546),

45 See ICA, cited above nt. 18, p. 62 ff. reporting that permissions requested before downloading an APP, which are labelled as "normal" by an APP store, do not even require an explicit consent by the users. See also
46 We think especially to the “right of access” (Article 15), under which users are entitled to know whether their data are being treated and for what goals; whether they have been transmitted to third parties (even in countries outside the EU) and what are the forms of protection foreseen in this latter case. Read in combination with Article 22, GDPR, users are also entitled to know what is the logic underlying automatic treatment of their own data.
“right to erasure” (Article 17\textsuperscript{47}), “restriction of processing” (Article 18\textsuperscript{48}), and especially the “right to object” (Article 21\textsuperscript{49}). How many know, that since May 25, 2018, under Article 3 of GDPR, European citizens are able to profit from said chances even if the data holder is situated in extra European countries, which are less favorable to the users’ rights of access and control? — a situation that is applicable e.g. to US data owners — whereas before the GDPR entered into force, in that case, the applicable law was the US’? (50)

It follows that by leveraging on the evoked well established biases, such as the no-reading, the “take-it-or-leave-it” (51), or the “free meal” heuristics, or, as recalled, the straight ignorance of the users about their rights of access and control over their data, platforms have a formidable leverage to exploit and, hence, increase their market power. (52)

\textsuperscript{47} Article 17 GDPR stipulates an obligation on the controller to the erasure of personal data that are no longer necessary in relation to the purposes for which they were collected or otherwise processed; for which the consent for processing has been withdrawn; or processing has been made unlawfully.

\textsuperscript{48} Under which users “have the right to obtain from the controller restriction of processing” if the accuracy of the personal data is contested; the processing is unlawful; the processing is contested for some reasons (e.g. defense of legal claims). In such cases treatment is suspended for the time necessary to establish if the contestation is well grounded.

\textsuperscript{49} The norm is formulated in a complex way, as it merges two very different hypothesis: that of data processing for direct marketing purposes (§2), and for reasons of public interest and exercise of public functions. In both cases the issue at stake is that data processing occurs without the user’s consent having been collected. The user can oppose to the treatment; however, while in case (a) treatment will cease definitely and data will be cancelled; in case (b) the treatment will only be suspended, and the controller will have to “demonstrate compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims”. If the controller fails to demonstrate that, the treatment shall cease.

\textsuperscript{50} See in particular Article 4.1, litt. a) and b), Directive 95/46 EC formerly applicable.


\textsuperscript{52} We will come later to another cognitive bias highlighted by Adam Candeub, of the
4. Exploitation of Digital Users’ Rights as a Third Hypothesis of Dominance-by-Exploitation

4.1 Antitrust profiles

The analyzed cognitive biases provide a formidable chance for abusing of users’ rights, i.e. in particular for profiling, exploiting commercially, and even ‘trading' personal data and profiles outside users’ knowledge and control, and thereby increasing market power.

That is a fundamental point of junction between protection of privacy and protection of competition. A risk thus emerges of what we would label “dominance-by-exploitation”. By the latter we mean that market power is increased by exploiting both the looser laws of certain non-European countries as concerns the processing of data of non-European citizens as well as the straight ignorance of the users about their rights and their cognitive bias. Further, once dominance is attained, it can be abused by either exploiting users or foreclosing competitors.

The question then arises as per whether that may amount to illicit conduct under competition law standards, and, in particular, whether it might be configured as an exploitative abuse. We are inclined to say ‘yes', both in the case of pre-existing dominant position and in that of the acquisition of such position thanks to said abuses of users’ rights and of exploiting their cognitive biases.

In its decision on the Facebook/WhatsApp merger of 2014, the Commission took the view that privacy concerns are outside the scope of competition rules, by stating that: “any privacy related concerns flowing from the increased concentration of data within the control of Facebook as a result of the Transaction do not fall within the scope of the EU competition law rules but within the scope of the EU data protection provisions”.

---

Michigan State University College of Law, in a 2014 article appeared on the J. of Law and Policy for the Information Society (9 ISLP 3, p. 407-434), and submitted as a factor of market power in the web economy (see after, under para. 5).

53 One should recall that Directive 95/46/EC was not applicable to non-European companies, no matter whether European citizens were affected or not.
protection rules”. (54) It, therefore, did not investigate whether the platform's market power could eventually be increased because of an augmented capability of data collection; nor did it investigate whether it was possible for Facebook to technically merge WhatsApp’s and its users’ digital identities and data.

Quite to the contrary, its focus remained fiercely the advertising side of the (two-sided) market, not the free one: “regardless of whether Facebook would start using WhatsApp user data to improve targeted advertising on Facebook’s social network, there would continue to be a large amount of internet user data that [would be] valuable for advertising purposes and that [would not be] within Facebook’s exclusive control.” (55) On this basis, it concluded that the merger would not give rise to serious competition doubts as regards the market for the provision of online advertising services.

When in 2017 it was established that Facebook was “at least negligent” in providing “incorrect or misleading information” to the Commission (§§ 89 and 90) during the merger investigation, in breach of Article 14 of Regulation 139/2004, the possibility for the platform to “automatically merge users’ profiles of different apps” was already technically feasible in 2014 (§ 80). What in particular was ascertained, was “the possibility of matching Facebook IDs automatically with WhatsApp users’ mobile phone numbers” (§ 77). To give an idea of the magnitude of such an automated profile merging, consider that the latter was, upon admission by the same defendant, already possible on Android OS. In 2014 the latter operating system was running on the majority of smartphones in the EEA, ranging between 60% and 75% (depending on the estimates).56

56 See European Commission dec. Case COMP/M.8228/2017, cited at the previous note, at § 80 and nt 59.
The conducts leading to such abuses are now more effectively checked by the GDPR Regulation no. 2016/679 entered into force on May 25, 2018. However, it must be emphasized that this Regulation, in particular, addresses profiling at several points (see in particular Arts 13, 14, 15, 22, and Recitals 60, 63, 72, 91; see also the Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 of the Data Protection Working Party [set up under art 29 of Directive 95/46 EC]). It is a quite advanced Regulation, indeed, although possibly hosting some loopholes (see e.g. the absence of the requirement of user’s consent for the processing of data provided by Art. 6.1., lit. f) under the [perhaps too] broad reference to the ‘legitimate interest pursued by the [data] controller or by a third party’; emphasis added).

Also, there is probably too much reliance on “transparency”, as Art. 12, in particular, establishes an obligation on data controllers to provide users with all information about their rights (of “access”, “erasure”, “restriction of processing” and “to object”) and how to exercise them. Albeit, it is widely known that providing consumers with more information through disclosure duties is most of the time useless because such information is highly technical and consumers lack the specific knowledge apt to understand it. Most decisions about how to dispose of one's own data are taken by impulse, without evaluating the real consequences of the (implicit) exchange. (57)

However, it must be emphasized that the GDPR is primarily European — in that it only protects European Citizens and users, even if the data holder operates outside Europe -- such as the US as well as China, India, Russia, just to mention the biggest, or even – in the near future – in the post-Brexit UK.

In this case, there are two ways to configure dominance-by-exploitation. The possibility to commercially exploit and trade data and digital profiles outside individual and SMEs’ control and awareness,

57 See ICA, cited above nt 18, p. 4.
leading to an increase of market power or its abuse, occur when the use of data:

(i) violates/exploits data protection rules and amounts to a possible abuse of dominant position under the meaning of Article 102 TFEU (or equivalent norms in other jurisdictions); or

(ii) violates/exploits different kinds of rules, such as contractual relationships between platforms and their business counterparts pertaining to data, also amounting to an abuse of dominant position.

In both cases, the competitive harm arising from dominance can be understood as one of increasing consumers’ switching costs. In particular, aiming to maintain their data advantage and/or prevent their rivals from attaining scale (foreclosure), dominant firms increase the switching costs for consumers, thus making it harder for consumers to leave to an alternative service (‘lock-in’ or ‘walled-garden’ effect). That is done, by obscure (to users) forms of data collection and treatment or by leveraging on their well-established biases and heuristics.

The EU case law recognizes that a dominant firm abuses its dominant position by gaining advantage from unlawful or deceptive acts that enable to exclude effective competition or exploit it. (58) In parallel with that, it should also be possible to question whether, once the firm has attained a platform scale and thus dominance, exploitation of users through loose contractual terms pertaining to data protection and collection, would be admissible. Also, the Bundeskartellamk acknowledges that "data obtained using unfairly broad terms and conditions can contribute" in determining a violation of competition rules. (59) And in its preliminary assessment of the Facebook abuse case,

---


it also acknowledges that Facebook’s terms of service “violate data protection rules to the disadvantage of its users”, which have no alternatives but to accept the whole package or not receiving the service. (60)

In the case of Facebook, already a violation of the processing of data has been established by the Commission with regard to users’ personal identities and phone numbers in breach of the undertakings for clearing the 2014 Facebook/WhatsApp merger. As anticipated, more undisclosed data transfers (allegedly) occurred between the platform and the major manufacturers of fixed and mobile devices, which eventually helped the tech firm gaining position in the market. Those agreements, in turn, seem to be in breach of a 2011 consent decree signed with the FTC.

Regarding the possibility that platforms exploit their dominant position to impose unfair and intransparent terms and conditions to their contractual counterparts (made especially of micro and SMEs – and the middle class at large 61), new evidence comes from recent EU legislative initiatives. In particular, it is acknowledged that platforms engage in a series of unfair contractual practices that would not be possible if they did not hold a dominant position. Such practices consist especially in: (i) sudden unilateral changes in terms and conditions without prior notice; (ii) the delisting of goods or services and the suspension of accounts without a clear statement of reasons; (iii) lack of transparency related to the ranking of goods and services and of the undertakings offering them; (iv) unclear conditions for access to, and use of, data collected by providers; (v) favouring of platforms' own services and use of the so-called most-favoured nation (MFN) clauses, which restrict the users' ability to offer more attractive conditions through other channels than the platforms’ services. (62)

60 Körber, cited at previous fn, p. 20.
62 See the Commission’s SWD cited above.
Clearly, whenever such conducts are put in place by dominant firms *because* of the exercise of their dominant position, and at the same time have a foreclosure effect, they would hardly escape a violation of Article 102 TFEU.

Now, a global phenomenon, based on a substantially global technological approach, affecting the entire data-driven economy should be regulated under a substantive harmonization. *And* one pro-user: first of all because this is intrinsically just *since the user is typically the weak party* in the relationship with a data holder that is a dominant platform. And also because the long-term general interest to the healthy development of the data-driven economy requires the citizens’ trust.

### 4.2 Socio-cultural and Informational profiles

On the connected information profiles, measures shall also tackle the problem of power that from the strictly economic plane overflows onto the societal level tout court, displaying a pervasive capacity to influence the very dynamics and pluralism of democracy ultimately. [Incidentally: shouldn’t we enlarge the perspective of the ‘user/consumer’ in a strictly individual microeconomic sense to that of *citizen* across-the-board: as such more interested in enjoying effective informational hence cultural hence political pluralism than in getting some price or non-price individual advantage advertised by ‘his master’s voice’?] That is achieved through the control of information, media and influential lobbying organizations: instruments of an overall dominance that cannot be assessed on purely micro-economic bases, as, i.e. Steven Solomon Davidoff of Berkely U. has persuasively argued).

Now — and this is a further point of junction between protection of privacy and protection of competition — that superpower, that societal hegemony of the web-titans (as The Economist recently labelled them) has grown so much as to facilitate same in conditioning — even in Europe, albeit especially in less pro-user regulated regions — the factual application of even severe rules — such as those of the soon in force EU GDPR, by influencing national legislators and authorities (less the Courts) towards the adoption of
milder standards of enforcement (i.a. policies of imposing fines only, not also structural remedies) and/or bureaucratically complicated, hence time and money consuming, ‘discouraging’ modes of access by the individual citizens to the knowledge of how their profiles are shaped, stored, used and for what purposes, resold and to whom? With what frequency? To do what? And, also, by discouraging institutions from engaging in broad national campaigns of citizens’ rights awareness — the fundamental ‘bottom-up’ long-term remedy to their cognitive biases.

Not to mention their lobbying power exercised to hinder or soften antitrust intervention and the ability to influence even financial markets: e.g. an announcement by Amazon's CEO to enter a new market can depress or the value of competing companies' shares (63).

5. Possible Remedies

Now, in a broader, meta-EU, ‘global’ perspective, the focus should be placed on two basic connected profiles.

One is that of which possible remedies are conceivable to contrast dominance-by-exploitation by both (A) regulatory and (B) antitrust tools.

The second profile is ultimately political and addresses the question of which measures —even beyond specific abusive conducts — might be put in place, without hindering the technological progress, in order to reduce the vast power (the superdominance, as Richard Wish would call it) that the GAFAM web oligarchs (Google, Apple, Facebook, Amazon, Microsoft) the five sisters of the web economy hold and exercise by controlling the big platforms where information circulates, as well as an increasing number and type of media (Zuckerberg himself famously declared that Facebook is also a media company).

63 Galloway, cited above, nt 61.
From a global perspective, reaching beyond Europe and European citizens, we believe that under the first profile:

(A) interventions of regulatory nature could be conceived of, stemming from structural market failures connected to the exchange of data in platforms (64):

a) Data protection rules on profiling, and their enforcement, should be strengthened, in non-European legislations, on the blueprint of the European GDPR: this in order to both protect all individual citizens from abuses of their privacy rights (thus also avoiding that easy data collection and thus profiling continue being a multiplicator for market power of digital platforms.

b) Also, interventions at the data level should be considered by

(i) allowing users to choose not to release their personal data, but to pay instead for the service or app they choose to use (65);

(ii) implementing transparency (disclosure) duties on profiling and reselling of personal profiles, along with the lines of the rights currently entrusted by the GDPR. Importantly, disclosure duties should be designed in a

64 ICA, cited above, nt 18, pp. 5 and 41 according to which, the exchange of data between users and platforms gives often rise to "structural market failures" because the huge investments on data collection about individuals made by digital firms do not internalize social costs, thus leading to overinvestment in information collection. Moreover, due to transaction costs and uncertainty as per the allocation of data property rights, it is highly probable that market forces do not lead to efficient outcomes. Thus, a situation occurs where the interests of those having an informational advantage (i.e. more technical knowledge about data) prevails. (information asymmetries, on which see above, para. 3)

cognitive-based fashion (66), that is: in a way to be salient and simplified enough to be quickly understood, so to increase awareness of both the "if" and "how much" of individual consent.

c) Data protection and Competition Authorities in cooperation should launch wide institutional awareness campaigns about users' rights and the exercise thereof. Indeed, as former European Data Protection Supervisor (67) said, a stricter collaboration between data protection supervisors and competition agencies in this regard is essential to “increase transparency” over profiling and re-combination of personal data. These awareness campaigns should be mandatorily hosted and re-launched by big platforms and social media, in the logic of public service that should permeate the whole sector of information to the public through the salient media.

All this, of course, should be done without prejudice of policies incentivizing of private players to develop ‘reactive’ algorithms, as the new apps that oppose to personalized prices for airplanes based on individual profiling. (68)

(B). Under the existing antitrust toolbox, both structural and behavioral remedies are conceivable — de lege lata.

At the behavioral level:

a) Making key data available to other firms, if a user so wishes, just as European banks must now do with clients’ account information,


67 See EDPS, Preliminary Opinion, Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy, March 2014, p. 8.

68 See, e.g. Hopper (https://www.hopper.com/), an APP that uses big data analytics to predict when flight prices are at their lowest points; it then notifies consumers that information to get the best deals on time.
so as to make it easier to switch between services. See here the model set forth by Art. 20 GDPR on personal “data portability”, which is applicable irrespective of the market power owned by digital firms.

b) We hesitate to support an ad-hoc obligation to share data, tailored upon the firms’ size (i.e. the bigger the firm or group, the more to share). Indeed, data-sharing is a typical “mode of production” of digital economy (and of platform positions). Instead, we would favor a duty to share more enabling technologies (e.g. AI-related) to avoid creating barriers hindering dynamic competition.

At the structural level: we would, first of all, acknowledge the incipient stage of the elaboration on this point, by both the European Authorities and scholars; and therefore the need that such type of measures does not discourage innovation-oriented investment policies.

That said, we think that:

a) mergers should be more robustly tackled whenever a deal/acquisition is likely to keep off future newcomers potentially. E.g., authorities should have prevented Facebook’s acquisition of Instagram and/or WhatsApp as well as Google’s acquisition of Waze (maker of navigation software), on the consideration that the mergers would have increased in a ‘conglomeral’ direction the two big ones already giant market power.

b) divesting: in the ‘liquid’, all-immaterial web environment, divestitures can hardly take place à la AT&T — there are no ‘antennas’ or ‘bridges’ or other physical infrastructures on which to work. Divesting is here possible on the corporate plane. E.g. by separating GAFAMs in as many entities as the type of function each entity of the group performs, by using the data collected when it provides its services. For instance, dividing Google could be divested in separate independent (and contendible) corporate entities: one for the search engine, one for the mail service, one for Youtube, one for the maps… Or, as per Facebook, it could be separated by dividing Instagram from WhatsApp, etc. Or, as to Amazon, it could be split in Amazon and Amazon Service, and so forth. Independent entities, would thus not operate under the strategic control of a single holding company.
We expect several objections. One is that divestiture has nothing to do with competition, since each of those companies would operate in separate markets. One can reply that a conglomereral structure, operating under a unitary strategic command is in condition to modify, by the joint effort of its components and the ‘deep pocket’ of the group, the level playing field of competition in the specific markets of its single units.

One could also reply by borrowing the argument set forth, as hinted, by Professor Candeub. In the web economy, competition is not ‘a click away’ as Chicagoan die-hards like Bork and Sidak affirm, suggesting that consumer loyalty is only due to superior offers. Online market behavior, Candeub argues (69), is based on the authoritative theories of Daniel Kahneman (70), that is: it is driven by behavioural tendencies related to habit and information costs. And the latter can make switching quite difficult for the average consumer. Searching the Internet is quite taxing cognitively, and consumers desire to decrease their cognitive costs. Human beings are more favorably disposed towards works and experiences to which they have already been exposed and easily understand. People tend to avoid strenuous mental efforts.

Thus, in short — we hastily summarize the argument, the article is complex and deserves deep attention — if Google search provides ways to lower these costs through convenient access to the different desired internet services, the switching costs — cognitive costs — that develop as Google use becomes habituated will significantly increase. And if these tendencies are magnified, as they often can be through network effects, it is at least possible that market power is increased, and anti-competitive results (de facto foreclosures) may follow. Candeub consequently argues that Google as a whole constitutes an

---

essential facility (or EF) (71): the EF consisting precisely in the links between its search box, YouTube, Google Books, maps, e-mail, and so on. These links provide overall easy, low cognitive-cost access to the web, particularly if Google’s use is reinforced, as hinted, by habits. Even more broadly, the EF is made of the portal or interface that Google provides to the web.

As to other (trivial) objections such as that the envisaged measures are vexatious and anti-business, one could reply with the above-hinted remark about the long-term convenience for the data-driven economy to increase the trust of citizens. In this connection we suggest reading of issue Jan 20, 2018, of a magazine bearing the cover story and cover title: ‘The new titans, and how to tame them’ (emphasis added). Which magazine? The Economist, hardly a subversive paper... (72)

Conclusively, if antitrust law and its application miss challenging the superdominance of those web Titans seriously, it will deserve John Galbraith’s famous sarcastic remark: that it shuts the stable door after the horse is gone, reserving its thunderbolts to parvenus…

May we should here recall that in introducing his Act, Senator Sherman famously emphasized that “if we will not endure a king as political power, we should not endure a king over the production, transportation and sale of any of the necessaries of life”. Were he living today, he would add information and data to the list. (73)

---

71 Candeub, cited nt 69, p. 410.