Personal Data Protection and Its Reflexes on the Data Broker Industry

Guilherme Birckan(✉), Moisés Lima Dutra, Douglas D. J. de Macedo, and Angel Freddy Godoy Viera

Universidade Federal de Santa Catarina, Florianópolis, Brazil
gbirckan@gmail.com

Abstract. Demonstrates the relationship between government and private interests in identifying people's profiles on the Internet. Describes the establishment and development of information aggregators and merchants, the data brokers. Discusses the boundaries of personal data commoditization, which in consequence wears away privacy and anonymity. Associates the inception of laws that mandate publicity to data breaches events, exposing the model, and ensuing debates on the need for further regulation. Presents the innovative generation of legislation created to govern a business that up until then operated free from public scrutiny. Introduces ideas to prevent the extinction of such a business model upon the shift to privacy and data protection.

Keywords: Digital identity · Internet privacy · Personal data · Data brokers

1 Introduction

Big Data, Cloud Computing, Cloudlets, Internet of Things, Data Brokers, all contemporary terms, buzzwords associated with what is being hyped in tech trends - except for the last one. Data brokers, although increasingly subject to debates, are as old as the net itself. After all, since information started being published online, there has always been a need for aggregators, as there always have been those who were interested in their products: compilations from sparse sources generating specific dossiers about something or someone.

There is power in identity, as according to Castells (2011), it "is people's source of meaning and experience". The most important link in the "who-what-where-when" tetrad, the unique identification of a being that is extracted from a data bulk (especially the Web) has been, for decades, the desire of states and corporations. Prins (2006) points out that "a look at our contemporary, data-based society reveals that information about people is essential for a variety of economically and socially useful and crucial purposes: education, taxation, social benefits, health care, crime detection and terrorism prevention, commerce and marketing, to name but a few".

On one side, ordinarily control-avid governments, armed with the national security argument; on the other, the refinement of targeted marketing, progressively individualized to declared and inferred tastes; in the middle, lies the citizen, the user, the target,

R. Mugnaini (Ed.): DIONE 2020, LNICST 319, pp. 103–117, 2020.
https://doi.org/10.1007/978-3-030-50072-6_9

fooled in that a mouse and a screen would reflect some degree of anonymity, while the truth is that every search, every click, every like, every post, are all cataloged, reunited, processed, and, not rarely (or maybe very frequently), sold to third parties.

Usual entry points for data acquirement, and at first sight somehow innocent, are the filling in of forms, commercial transactions, internet searches, use of social network platforms, webmail, loyalty and discount programs - including websites, retail, banks, drug stores and health plans, among many other interactions. Info yield can be carried in either an active way, when the user consciously inputs or allows such capture, or passive, when data is collected without actual acknowledgment - for instance when conversations or images are recorded, when messages are read by algorithms, or GPS history is logged.

The goal of this paper is to recognize the new generation of laws that have been created worldwide, and that have in their cores the establishment of principles, limits, and responsibilities to those that produce and consume data of individuals. Initially, concepts related to the data business as well as a brief history of such industry will be introduced. Next, some of the events related to the leak of personal data and that added to discussing the need for regulations will be illustrated. At last, legislation that was recently enacted in the United States, Europe and Brazil will be acquainted, and their effects on the data industry will be debated.

2 Data Brokers

Data brokers are companies that collect information, including personal information about consumers, from a wide variety of sources for the purpose of reselling such information to their customers for various purposes, including verifying an individual's identity, differentiating records, marketing products, and preventing financial fraud (Federal Trade Commission 2012). For the extraction (and presumption) of knowledge, statistical algorithms are used, and, nowadays, Horvitz and Mulligan (2015) add that machine learning techniques are also used, which can facilitate making the leaps across informational and social contexts, generating inferences.

Big Data, as described by De Mauro et al. (2016), "is the information asset characterized by such a high volume, velocity, and variety to require specific technology and analytical methods for its transformation into value". This definition was picked for this paper for its objectivity and simplicity, but considering the ubiquity of the noun, such is not uncontested, as Ward and Baker (2013) explain: "owing to a shared origin between academia, industry and the media there is no single unified definition, and various stakeholders provide diverse and often contradictory definitions".

The abundant volume of information that people generate every second didn't take long to have its potential recognized, becoming, beyond a commodity, a whole new specialized business. As Sevignani (2013) explains, "commodification is the process of making things exchangeable on markets, either actually and/or discursively by framing things as if they were exchangeable". Roderick (2014) adds that "the growth of companies […] has facilitated a shift in attention from production-oriented to marketing-oriented strategies, allowing companies to tap into and encourage (ir)rational purchase behavior". More than random or spontaneous data, especially nourished by the massive scale of the social networks, Big Data is also constructed from individuals' data, and that is where its real value resides.

The inception of the data broker industry brought, at its foundation, an extensive list of attracted parties. Regular customers range from banking institutions and financing agencies, seeking risk mitigation and fraud detection, to service providers from niche markets and online retailers, who are interested in consumption profiles, to politicians and candidates, who seek to know - and to influence - their audiences, and also security and intelligence agencies, interested in strengthening their investigations and predictions. Mosco and Wasko (1988) explain the essence of what is happening: "new technology makes it possible to measure and monitor more and more of our electronic communication and information activities. Business and government see this potential as a major instrument to increase profit and control. The result is a pay-per society". Figure 1 illustrates an ordinary flow of information to/from a data broker, describing commonly used sources for capturing data, as well as other public and private actors who participate in the ecosystem.

Although governments usually possess robust databases, eventually they also end up hiring data brokers' services, as according to Stevens (2001), "private companies maintain and organize personal information on individuals in a manner that may not be legally available to government actors". As an example, there is the United States Privacy Act[1], which "establishes a code of fair information practices that governs the collection, maintenance, use, and dissemination of information about individuals that is maintained in systems of records by federal agencies".

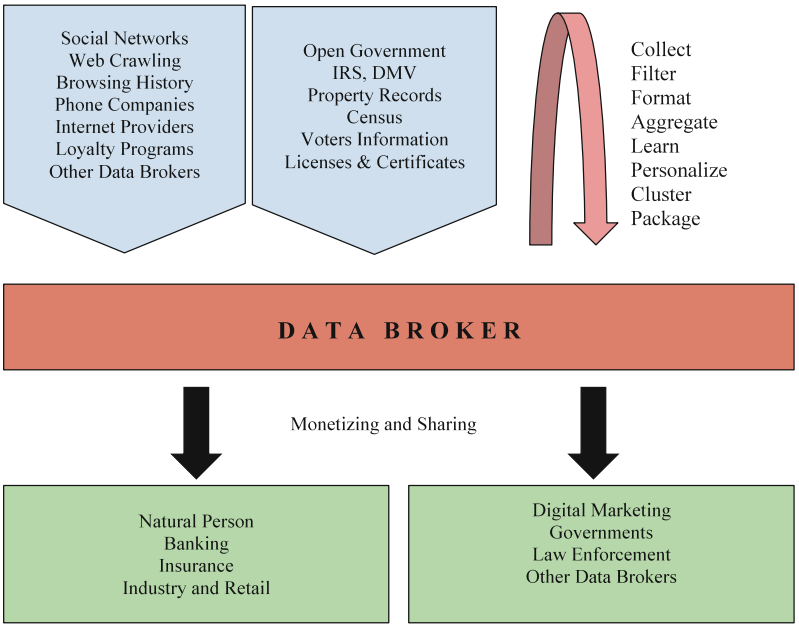

Fig. 1. Illustration of an ordinary flow of information from/to a data broker, inspired by the work of Otto et al. (2007).

[1] https://www.justice.gov/opcl/privacy-act-1974.

The comprehension of an individual's decision process through the employment of statistical models allows the establishment of behavioral patterns; additionally, it also grants the development of anticipation on trends (Ostrowski 2013; Ngai et al. 2009). Examples include habit changes, fluctuation on demands, and interest shifts for specific goods. This predictive value complements the intrinsic importance of static data: the greater the collection of entries on someone, the preciser the classification models will be, thus justifying the rush on digital mining.

Knowing information about people makes it possible to cluster them, which means assemble or label them when they share similar characteristics - or according to requirements and attributes pertinent to whoever is interested, from socioeconomic profiles to consuming inclinations. This is a capability inhabited by a latent ethical impasse, as it opens the possibility of the usage of variables that are not only merely demographic, but that may imply questionable context. Features that nowadays are not seen as politically correct involve race, religion, gender, age, and income, among others – and in some legislations could also be a crime. Therefore, linking digital profiles to automated decision-making algorithms may (inadvertently or purposely) lead to discriminatory results, as pointed in the Big Data and Privacy report[2] prepared for The US President's Council of Advisors on Science and Technology in 2014. The report examined the nature of current technologies for managing and analyzing Big Data and for preserving privacy, it considered how those technologies are evolving, and it explained what the technological capabilities and trends imply for the design and enforcement of public policy intended to protect privacy in Big Data contexts. Among its conclusions is the recommendation that policy should focus primarily on whether specific uses of information about people affect privacy adversely. It also recommends that policy focus on outcomes, on the "what" rather than the "how," to avoid becoming obsolete as technology advances.

3 Data Privacy

Although a growing market and virtually multibillionaire, the dissemination of the data broker business model did not occur without questioning. The debates orbit around recurrent subjects, among which, three are commonplace: transparency for when data is captured, loss of control over one's anonymity, and the sharing of the profits.

On the (lack of) transparency subject, Reyman (2013) describes the reality that "terms-of-use policies that describe data collection and use are required by law, but these are lengthy and difficult to understand when read at all" and that data is often obtained on social web technology trade-offs from "tacit agreements that users enter into, and a set of unspoken assumptions that govern who owns what is created and how it circulates." Gangadharan (2017) added that, during research, "marginal Internet users ignored privacy policies or terms of service agreements that they encountered", and "when signing up for email, and despite instructors' advice to carefully review user agreements, students clicked through or past privacy policies and terms of service in order to complete the registration, suggesting these notification mechanisms functioned as meaningless accessories to the new learner's Internet experience".

[2] https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf.

The next common question refers to where do the data end up after all (?), and who has access to it (?), key issues on the argument of anonymity control. In a non-regulated system, information can be sold and distributed without acknowledgments or even accountability of the transactions. In that prospect, Rachels (1985) explains that the value of privacy is "based on the idea that there is a close connection between our ability to control who has access to us and to information about us, and our ability to create and maintain different sorts of social relationships with different people". Roessler and Mokrosinska (2013) add that "the control and regulation of informational privacy should be viewed not only under the perspective of individual rights, but also as being necessary for social interactions themselves, and therefore as relevant to the integration of society".

Another controversy that is consistent concerns the earnings from third-party information: if companies in such business make extraordinary profits with data that are essentially generated by people, where are my paychecks? In that direction, Malgieri and Custers (2018) describe how "personal data of individuals represent monetary value in the data-driven economy and are often considered a counter-performance for 'free' digital services or discounts for online products and services", and point out that "individuals do not seem to be fully aware of the monetary value of their personal data and tend to underestimate their economic power within the data-driven economy and to passively succumb to the propertization of their digital identity".

In a nutshell, the essential aspect of the data broker industry, as emphasized by Crain (2018), is the asymmetrical loss of privacy: "people are opened up to increasingly extensive forms of monitoring, while the institutions doing the monitoring and the information they collect remain hidden from view. […] Privacy asymmetry as a descriptive category is especially salient for the data broker industry, which has long operated without public awareness or direct regulatory oversight. The privacy of those under watch is undermined, while the watchers themselves operate with substantial freedom from scrutiny".

As technology advanced and the Internet's popularity escalated, there was a great increase in the availability of digital knowledge, as much in the Web as in private databases. Despite the fact that the flourishing of online information might also have dilated the offering of raw material and the number of players that use them, data aggregators have always existed. The catalysts of public objection and the beginning of the model exposure, after decades of progressive exploring, were events known as data breaches, as we will see next.

4 The Exposure of the Data Brokerage Industry

A data security breach occurs when there is a loss or theft of, or other unauthorized access to, sensitive personally identifiable information that could result in the potential compromise of the confidentiality or integrity of data (Stevens 2012). Legislation that addresses such cases usually requires the events to be made public, and both the potentially affected individuals and regulatory agencies to be informed. The obligation to make the facts known is broad, reaching not only data brokers but any private, non-profit or public organization, regardless of their area (health, education, insurance, finance, etc.).

Stevens (2012) explains that security breach notification laws generally follow a similar framework and can be categorized into several standard elements: (1) delineating who must comply with the law; (2) defining the terms "personal information" and "breach of security"; (3) establishing the elements of harm that must occur, if any, for notice to be triggered; (4) adopting requirements for notice; (5) creating exemptions and safe harbors; (6) clarifying preemption and relationships to other federal laws; and (7) creating penalties, enforcement authorities, and remedies.

The significance of the expanding expenses in cybersecurity, with the added intent of also preventing - or minimizing - breaches, can be observed in Fig. 2, where it is illustrated the annual spending, in the United States, in that area.

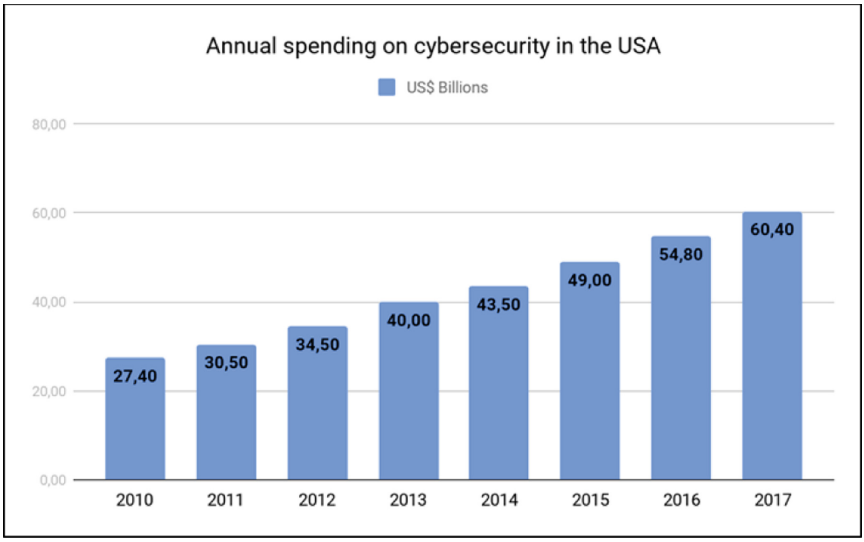

Fig. 2. Annual spending on cybersecurity in the USA (data source: statista.com).

In turn, from the annual records of data breaches in American organizations for the decade 2008 to 2017 (Fig. 3), despite steady expenses in cybersecurity, employee training and expanding regulations concerning individual data maintenance, incidences have been following a stable pattern. Even though the statistics display apparent regularity on the annual number of events illustrated in Fig. 3, on the other hand, the volume of compromised records has been following a rising trend, as can be seen in Fig. 4. It is worth noting that, according to the research, one record corresponds to one data entry, but not necessarily one single individual, as in a computational system, the same person may own several user accounts (for instance, using different email addresses). Considering this ascending scenario, one possible explanation could be the increase in database sizes, proportional to the popularizing of the social networks and the employment of Big Data technologies for capturing information.

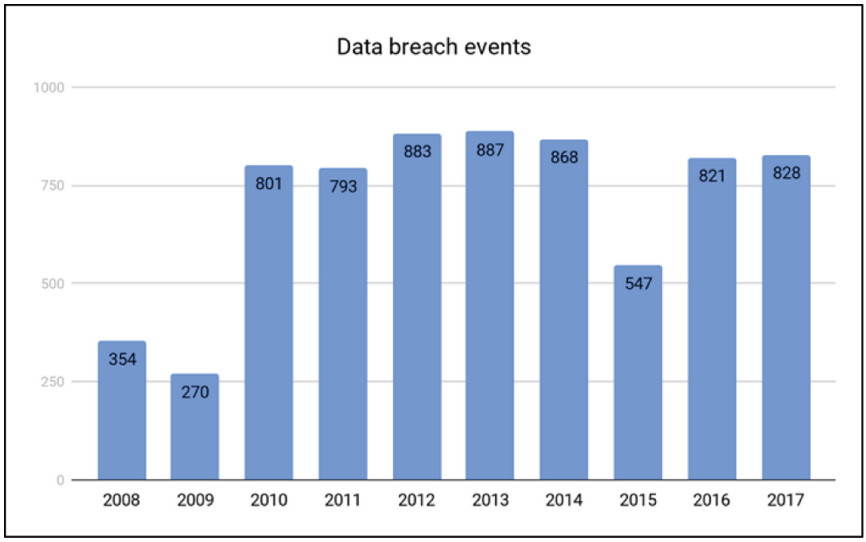

Fig. 3. Data breach events in American organizations in the decade 2008–2017 (data source: privacyrights.org).

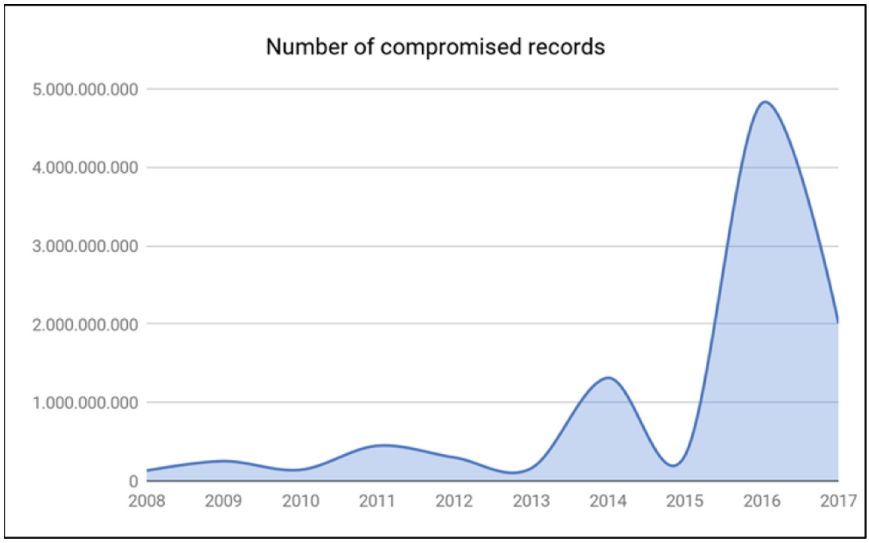

Fig. 4. Progress of the number of compromised records, annually, in data breach events (data source: privacyrights.org).

Data breaches can be categorized according to their causes or origins, which are not limited to cyber-attacks from hackers and malware (although most of the incidents are based on those, granted Fig. 5). They are also considered the cases of unintended disclosure (sensitive information posted publicly, mishandled or sent to the wrong party),

physical loss (paper documents or portable devices that are lost, discarded or stolen), insider (someone with legitimate access intentionally breaches information), fraudulent transactions involving debit and credit cards, and finally the unknown cases. Unauthorized access to data causes direct losses due to financial fraud, identity theft, and industrial espionage, besides indirect losses such as reputation and asset depreciation.

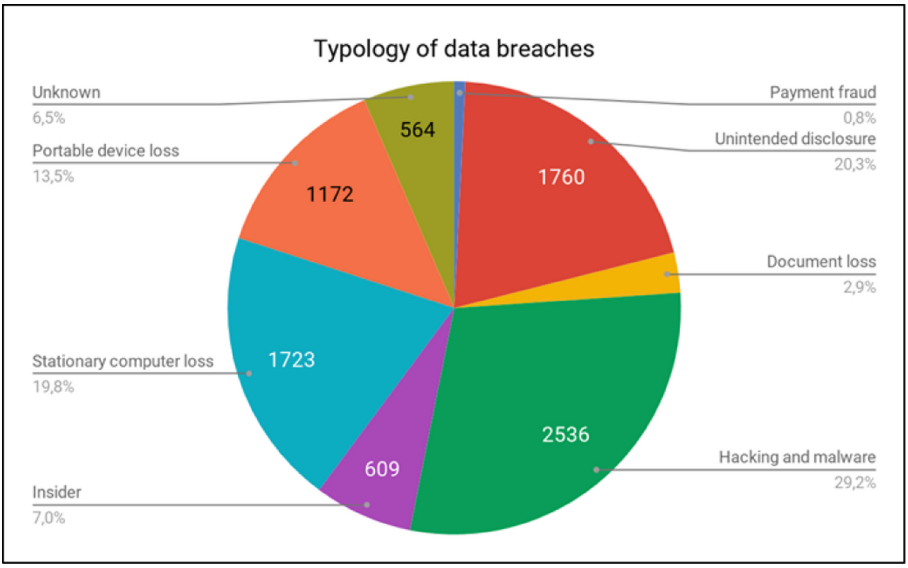

Fig. 5. Typology of data breach events between 2008 e 2017 (data source: privacyrights.org).

Identity fraud, according to Gordon et al. (2007), may be defined as the misuse of personal or financial identifiers for personal gain or to facilitate other criminal activity. Such gains may be obtained from online shopping scams, usage of stolen cards, or controlling and spending over someone else's accounts. Information leaks are often the fuel and/or the beginning of cybercrimes, and the increasing statistics of such frauds, added to the growing cases of data breaches, ultimately brought the attention to a market that up until then operated quietly.

Those who suffer from crimes of identity theft or misuse of personal data are customarily left, often with little or no assistance, with the necessary bureaucracy to reestablish their names and losses. Side with the clamor of the victims, there has been some support from nonprofits and research centers that turned their attention to personal data protection. Among organizations that became distinguished on advocating for more privacy, preeminent cases are the EPIC - Electronic Privacy Information Center[3], and the Privacy International[4], which are both plaintiffs on several civil lawsuits about alleged privacy abuse from tech companies.

[3] http://epic.org.
[4] http://privacyinternational.org.

One can adduce that there has been a first era of exploring and preying on personal data on the Web, due to the lack of regulations and auditing. Recent movements, with the enactment of laws specifically designed with the focus on information privacy, point towards a paradigm shift, where there's empowerment to the users while managing their data. In the next section, examples of the legal innovations that sustain this transition will be presented.

5 Personal Data Usage Regulations

It has been observed that technology is likely to evolve faster than the legislative processes. Therefore, the classical narrative applied to the approach of a new problem is to try to address it with existing older laws, up until the creation and establishment of specific regulations. In Brazil, recent examples of this reality are the laws 12.965/2014[5] (known as Marco Civil) and 13.640/2018[6], which regulates the paid transport of passengers by private drivers (like Uber and Cabify, business models created by portable apps). Before the Marco Civil, duties and rights of users and internet providers were disciplined by the Civil Code, Penal Code, Consumer Defense Code and the Constitution (all of them older than the Internet itself).

Pires and Cauchie (2011) state that a legislative change is a deviation chosen by the political system to create an innovation in the law structure, so it can be said, both theoretically and empirically, that it performs an innovative result. Once that innovation is a political operation, but belongs to the legislation (and not to the systems of thought), it produces, at the same time, but in another strand, a special impact on the normal and cognitive arrangements of the law system.

The motif of personal data privacy started to get more attention upon scandals involving information security leaks, obligated to be disclosed by laws that came at the beginning of the third millennium. One example was the enactment of a pioneer Californian bill[7] in 2002, which requires "a state agency, or a person or business that conducts business in California, that owns or licenses computerized data that includes personal information, as defined, to disclose in specified ways, any breach of the security of the data, as defined, to any resident of California whose unencrypted personal information was, or is reasonably believed to have been, acquired by an unauthorized person".

In light of the spreading number of exposure reports (displayed in Fig. 4) and of identity theft crimes, consequently, the model of data acquisition commenced being challenged. The next generation of laws, after a maturing cycle that took ten to fifteen years, is presenting itself much more rigorous when it comes to managing personal data. The main example was the enactment, in 2018, of the European General Data Protection Regulation (GDPR)[8], which affects companies that conduct businesses in the European

[5] http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm.

[6] http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Lei/L13640.htm.

[7] http://www.leginfo.ca.gov/pub/01-02/bill/sen/sb_1351-1400/sb_1386_bill_20020926_chaptered.pdf.

[8] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX%3A32016R0679.

Union, regardless of where they're headquartered. The major changes introduced by it were synthesized in Fig. 6.

Clear language	
Before	After
Often businesses explain their privacy policies in lengthy and complicated terms	Privacy policies will have to be written in a clear, straightforward language
Consent from user	
Before	After
Businesses sometimes assume that the user's silence means to consent to data processing, or they hide a request for consent in long, legalistic, terms and conditions - that nobody reads	The user will need to give affirmative consent before his/her data can be used by a business. Silence is no consent
More transparency	
Before	After
The user might not be informed when his/her data is transferred outside the EU	Businesses will need to clearly inform the user about such transfers
Sometimes businesses collect and process personal data for different purposes than for the reason initially announced without informing the user about it	Businesses will be able to collect and process data only for a well-defined purpose. They will have to inform the user about new purposes for processing
Businesses use algorithms to make decisions about the user based on his/her personal data (e.g. when applying for a loan); the user is often unaware of this	Businesses will have to inform the user whether the decision is automated and give him/her a possibility to contest it

Fig. 6. Key points established by the European General Data Protection Regulation (Source: https://ec.europa.eu/commission/sites/beta-political/files/data-protection-factsheet-changes_en.pdf)

Stronger rights	
Before	After
Often businesses do not inform users when there is a data breach, for instance when the data is stolen	Businesses will have to inform users without delay in case of a harmful data breach
Often the user cannot take his/her data from a business and move it to another competing service	The user will be able to move his/her data, for instance to another social media platform
It can be difficult for the user to get a copy of the data businesses keep about him/her	The user will have the right to access and get a copy of his/her data, a business has on him/her
It may be difficult for a user to have his/her data deleted	Users will have a clearly defined "right to be forgotten" (right to erasure), with clear safeguards
Stronger enforcement	
Before	After
Data protection authorities have limited means and powers to cooperate	The European Data Protection Board grouping all 28 data protection authorities, will have the powers to provide guidance and interpretation and adopt binding decisions in case several EU countries are concerned by the same case
Authorities have no or limited fines at their disposal in case a business violates the rules	The 28 data protection authorities will have harmonized powers and will be able to impose fines to businesses up to 20 million EUR or 4% of a company's worldwide turnover

Fig. 6. (*continued*)

Brazil also followed the international movement and introduced an update to the Marco Civil law of 2014 by the enactment of law 13.709/2018[9], which addresses personal data protection specifically. In that bill, ten cardinal principles are instituted, in consonance with values that are observed in the European version, and which are described in Fig. 7.

[9] http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm.

Principle	Description
purpose	data must be handled with legitimate, specific and explicit purposes, the user has to be informed, and data should not be processed later on for different reasons than it was initially acquired for
suitability	data must be handled in a way that is compatible with the goals informed to the user
necessity	data handling must be restricted to the minimum necessary to fulfill the purpose it was acquired for
free access	user's access must be provided in a free and facilitated way to information about how and how long their data will be handled
data quality	users must be assured about the correctness, clearness, relevance, and currentness of their data
transparency	users must be given clear, precise and easily accessible information about their data usage and handlers
security	establishment of administrative and technical rules to protect personal data from unauthorized access and from incidental situations such as destruction, loss, alteration, or disclosure
prevention	following of means to prevent damage due to data mishandling
non-discrimination	data cannot be used for abusive, illicit or discriminatory means
accountability	handlers must prove effective actions to obey and enforce such principles

Fig. 7. Principles that must be obeyed on personal data handling (source: Brazilian law 13.709/2018).

That law still establishes, as a requirement for data acquisition and processing, the user's consent, plus his/her right to access, manage, correct, and eliminate the data. Among the penalties for transgressions, there's a fine that can reach up to US$ 12 million. The milestone from the Californian bill that brought up attention to the data exploration business was also not forgotten: any security incident or breach that might lead to risk or damage must be disclosed to the corresponding individuals and the authorities.

Complementing the law 13.709/2018, in July 2019 was enacted the law 13.853/2019, which created the National Data Protection Authority, agency that has technical and decisory autonomy, is bound to the President's Office, and who has the responsibility (amidst others) of watching over personal data protection, overseeing that the rules are followed properly, and applying penalties.

Yet, regardless of Brazilian efforts, it is fair to point that the country was not the first in Latin America to implement those actions; Uruguay had a personal data protection law since 2008 (Ley 18.331[10]), while Argentina has had such legal framework since 2000, based on law 25.326[11]. The importance of those mechanisms was recognized by the European Commission, which regarded both countries "as providing an adequate level of protection for personal data as referred to in Directive 95/46/EC", per decisions 2003/490/EC[12] and 2012/484/EU[13]. Paraguay, in turn, has a bill[14] that was filed in March 2019 ("Proyecto de Ley de Protección de Datos Personales") and is still being discussed on their senate. Over the next section, there will be presented some ideas about adapting the data brokerage business to the new legal scenario described.

6 Proposals and Perspectives

For decades, players known as data brokers operated in a market with little to no regulation, where transactions between corporations (and even governments) were conducted without restrictions and free from public scrutiny. In that context, digitally, pretty much everything was possible: from capturing to buying and selling to sharing information, including that which was mined and statistically inferred (such as the clustering of profiles or the prediction of trends).

In a first moment, laws were crafted to make public those events known as data breaches, where it became mandatory that individuals were notified when their information had been accessed by unauthorized third parties, and which eventually brought attention to a whole market of personal data. The ensuing annual reports, containing a growing number of compromised records - despite the spending on cybersecurity, added to the advocates for privacy, and the tension from organizations that act on behalf of data protection, eventually contributed to the recent advent of a new generation of harsher regulations, already in a global scale.

Modern and important milestones were the European General Data Protection Regulation and, in Brazil, law 13.709, both from 2018. Such novel laws, whose restraints directly affect enterprises in the data brokerage business, suggest a possible downfall to the age where personal information was treated as a commodity. Nonetheless, adaptations to a new paradigm, more focused on privacy and data protection, are feasible, respecting the users and their individual authority on controlling who they are (their digital identities) and what they produce (their generated data).

Values and principles that guided the creation of data protection laws comprise (preceding) consent from the user, transparency, and purpose, premises that must be complied with and also considered while designing the business models. Shifting the spectrum from using third party data as sheer input, and bringing the original sources closer - the users - as partners and suppliers, it is conceivable to depict the continuity of several segments. With a certain level of anonymity or voluntary exposure, products

[10] https://www.impo.com.uy/bases/leyes/18331-2008.
[11] http://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/norma.htm.
[12] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32003D0490.
[13] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32012D0484.
[14] http://silpy.congreso.gov.py/expediente/115707.

such as behavior prediction and consumer profiling, or a wide range of classifiers, might still be appealing and functional.

From a collaborative perspective, new proposals arose, such as the policy framework for user data sharing by Iyilade and Vassileva (2013), based on the idea of a market. In that concept, applications can "offer and negotiate user data sharing with other applications according to an explicit user-editable and negotiable privacy policy that defines the purpose, type of data, retention period and price".

Malgieri and Custers (2018) investigated different models for quantifying the value of personal data, analyzing whether consumers/users should have a right to know the value of their data; the authors also discussed active choice models, in which users are offered the option to pay for online services, either with their personal data or with money. The conclusion, however, was that these models are incompatible with current data protection laws.

Tona et al. (2018) presented "a conceptual design for an artifact that will raise awareness amongst individuals about Big Data ethical issues and help to restore the power balance between individuals and organizations". Their proposal was constructed upon five dimensions, derived from the European GDPR, which are consent, the right to be forgotten, the right to access, data portability and data circulation. All those pillars are arranged over a foundation that would allow several collaborative interactions such as replying, commenting, reviewing, rating and tagging data.

Observing the ubiquity of mobile smartphone usage and the ensuing massive generation of data from those devices - locations, movements, images, video, text, and even health data, which is ordinarily uploaded to content-service providers, Mun et al. (2010) presented a privacy architecture named Personal Data Vaults. In their proposal, individuals retain ownership of their data, which can be reviewed and filtered before being shared, exploring three mechanisms for managing data policies: granular access control lists, trace-audit and a rule recommender, which provides a high-level interface for setting sharing policies.

It can be concluded that, in spite of the strictness of the novel privacy laws, albeit fresh and assigning tech companies to an adjustment cycle, there are several studies and initiatives towards creating tools (frameworks, models and architectures) that invest the users with more power to control their personal information. Such possibilities enforce a pattern shift away from the commoditization of data by the brokers - instead of the lone extinction of their model. Opportunities, therefore, might be in the effective deployment of collaborative platforms and products, which should have their primary focus on the value that has been so emphasized: the transparency.

References

Castells, M.: The Power of Identity, vol. 14. Wiley, New York (2011)

Prins, C.: Property and privacy: European perspectives and the commodification of our identity. Inf. Law Ser. **16**, 223–257 (2006)

Federal Trade Commission: Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers. Federal Trade Commission, Washington, DC (2012)

Horvitz, E., Mulligan, D.: Data, privacy, and the greater good. Science **349**(6245), 253–255 (2015)

De Mauro, A., Greco, M., Grimaldi, M.: A formal definition of Big Data based on its essential features. Libr. Rev. **65**(3), 122–135 (2016)
Ward, J.S., Barker, A.: Undefined by data: a survey of big data definitions (2013). arXiv preprint arXiv:1309.5821
Sevignani, S.: The commodification of privacy on the Internet. Sci. Public Policy **40**(6), 733–739 (2013)
Roderick, L.: Discipline and power in the digital age: the case of the US consumer data broker industry. Crit. Sociol. **40**(5), 729–746 (2014)
Mosco, V., Wasko, J. (eds.): The Political Economy of Information. University of Wisconsin Press, Madison (1988)
Stevens, G.M.: Data brokers: background and industry overview. Wall Street J. **6**(5), 552a (2001)
Otto, P.N., Antón, A.I., Baumer, D.L.: The ChoicePoint dilemma: how data brokers should handle the privacy of personal information. IEEE Secur. Priv. **5**(5), 15–23 (2007)
Ostrowski, D.A.: Identification of trends in consumer behavior through social media. In: 17th World Multi-conference on Systemics, Cybernetics and Informatics, WMSCI 2013, Orlando, Florida, pp. 9–12, July 2013
Ngai, E.W., Xiu, L., Chau, D.C.: Application of data mining techniques in customer relationship management: a literature review and classification. Expert Syst. Appl. **36**(2), 2592–2602 (2009)
Gordon, G.R., Rebovich, D.D.J., Choo, K.S.: Identity fraud trends and patterns. Center for Identity Management and Information Protection, Utica College (2007)
Reyman, J.: User data on the social web: authorship, agency, and appropriation. Coll. Engl. **75**(5), 513–533 (2013)
Gangadharan, S.P.: The downside of digital inclusion: expectations and experiences of privacy and surveillance among marginal Internet users. New Media Soc. **19**(4), 597–615 (2017)
Rachels, J.: Why privacy is important. In: Ethical Issues in the Use of Computers, pp. 194–200 (1985)
Roessler, B., Mokrosinska, D.: Privacy and social interaction. Philos. Soc. Crit. **39**(8), 771–791 (2013)
Malgieri, G., Custers, B.: Pricing privacy – the right to know the value of your personal data. Comput. Law Secur. Rev. **34**(2), 289–303 (2018)
Crain, M.: The limits of transparency: data brokers and commodification. New Media Soc. **20**(1), 88–104 (2018)
Stevens, G.: Data Security Breach Notification Laws. Congressional Research Service (2012)
Pires, A.P., Cauchie, J.F.: Um caso de inovação "acidental" em matéria de penas: a lei brasileira de drogas. Revista Direito GV **7**(1), 299–329 (2011)
Iyilade, J., Vassileva, J.: A framework for privacy-aware user data trading. In: Carberry, S., Weibelzahl, S., Micarelli, A., Semeraro, G. (eds.) UMAP 2013. LNCS, vol. 7899, pp. 310–317. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38844-6_28
Tona, O., et al.: Towards ethical big data artifacts: a conceptual design. In: Proceedings of the 51st Hawaii International Conference on System Sciences, January 2018
Mun, M., et al.: Personal data vaults: a locus of control for personal data streams. In: Proceedings of the 6th International Conference, p. 17. ACM, November 2010