The value of our personal data in the Big Data and the Internet of all Things Era

Anahiby Becerril

Doctor in Law and Globalization, Centro de Investigación e Innovación en Tecnologías de la Información y la Comunicación (INFOTEC). anahiby.becerril@infotec.mx

KEYWORD

ABSTRACT

Big Data; IoT; Personal; Value

The fourth industrial revolution, in which we transit, is preceded by a digital economy based fundamentally on information. And although estimates have been made on whether the data moving in the digital economy flow are from sensors or machines or come from each of us, one thing is certain, all information has a monetary value. Within this flow of information are our personal data. Every moment that we use an electronic device we leave behind vestiges of our life, which are collected by the machines to generate value to the companies. In this way our information is subject to market rules, supply and demand. We have become intangible beings, mercantilized, giving our bodies of data to the science, innovation and technological development. With the use of technologies such as Big Data and IoT, more information is less better. The speed with which our information is collected and treated as well as commercialized is undermining confidence in the digital market. Concern about the misuse of our personal data, or about the information we know about us, raises fundamental questions about privacy, ownership of information and human rights. The question of who should benefit from products and services based on digital data (generated by users) are the main uncertainties that shape the digital market opportunities. But, how can we measure trust in companies if we are not aware of the value our personal information has in the digital economy? What is the value that, as headlines, we give to our personal data? Perhaps therein lies the problem of allowing others to have power over our information. To promote the digital market and trust in it, we need to know the value of our information. Becoming aware of the fact and consequences of collecting our data, as well as its monetary value and importance in the digital economy, is the first beneficial step towards the empowerment of our information.

1. Introduction

We are currently moving into the fourth industrial revolution where mobile communications, social networks and sensors are blurring the boundaries between people, the Internet and the physical world (Marcus, 2015). This new revolution is preceded by the digital economy where markets leave space for networks and access constitutes a property (Rifkin, 2013, p. 14).
Companies in the technological sector lead the digital economy. All of them handle large volumes of information. For Pentland (2013) the digital economy allows us to study billions of exchanges made by individuals (p. 34) where people daily make transactions with ideas, money, goods or comments. In this sense, the greatest opportunity for growth revolves around the extraordinary volume, variety and timeliness of available digital data. The Internet, mobile phones, integrated sensors to equipment and other electronic devices generate data daily (Brynjolfsson & McAfee, 2016, p. 123). Every minute that we use some electronic device we are leaving traces of our life. Electronic devices become collectors of the information that we all generate. And all this information allows companies to gain competitive advantage and profits. From this perspective, information is presented as a key element of power. This is subject to market rules, supply and demand, in addition to being consumed, stored, processed (Araujo Carranza, 2009, p. 69) and exchanged. Our personal data have been immersed in this commodification of information.

We currently live in a tsunami of data (Bauman & Lyon, 2015, p. 31). Everything is data (Simon, 2013, p. 29) or it will be. As individuals, what defines us is no longer our person, but our value in data. We have been dematerified and we have become the digital “zeros” and “ones”2, irrigated in companies’ databases and it is through these that we demonstrate our existence.

The essence of social networking services is in exposing our privacy and sharing personal information. As users of Facebook, Twitter, Instagram, Pinterest and other types of content online, we not only consume this “free” content, but we have become the intangible asset of companies. We have become marvelous beings; our information, ideas, experiences, habits and feelings are accessible on the Internet. Facebook and Google, do not sell anything to Internet users, instead they sell billions of users to advertisers, which is why the wealth of ineludible companies is their number of connected/users (Ramonet, 2015, p.). As users of social networking and technology services, we have been commodified, and the voluntary delivery of our personal information is the price to pay not to suffer from social exclusion.

For the United Nations (2014, p. 7) and the World Bank (2016, p. 20), the extent to which we as consumers or users of electronic media really know what data we are sharing, how and with Who, as well as the use that will be made of the information, and in some cases, the companies do not communicate it. The World Economic Forum (2014, page 3) has recognized the existence of an asymmetry of power between institutions and individuals created by an imbalance in the amount of information about people and the lack of knowledge and capacity of the same to control the use of the information.

Internet is largely free service, and the exchange of personal data, is the price we pay (Ehrenberg, 2014) for its access. And while some people are willing to pay with this information, others consider that the use of this information, often without consent, may constitute a violation of privacy.

The speed with which our information is collected and treated as well as commercialized is undermining confidence in the digital market. Concern about the misuse of our personal data, or about the information we know about us, raises fundamental questions about privacy, ownership of information, global governance3 and human rights. The question of who should benefit from products and services based on digital data (generated by users) are the main uncertainties that shape the digital market opportunities (Marcus, 2015). But how can we measure trust in companies if we are not aware of the value our personal information has in the digital economy? What is the value that, as headlines, we give to our personal data? Perhaps therein lies the problem of allowing others to have power over our information.

1 Among the companies with the highest value of the sector are: Apple, value 586 billion US dollars; Facebook, value 314.8 billion; Amazon: 292.6 trillion; Google: 82.5 billion.

2 Referring to the binary system consisting of “0” and “1”. System used by computers to store information.

3 That is why we also aim for a governance of personal data as a response to the existing asymmetries in the market and data. Through the creation of technological tools, based on users that promote co-responsibility, in the collection and processing of information, between individuals, companies and the government. In this way there may be a coordination that implies benefits for all parties in the use and market of information.
2. Big Data and The Internet of things, in the use of our data

Characterized as “the new oil” and “a new class of assets” (Kuneva, 2009, p. 1), personal data are at the epicenter of the digital economy, as determined by Big Data and in the near future, IoT (acronym for Internet of Things). And if digitization and therefore data generation is considered “new oil”, the context through which this information is generated is now the new oil field, with industries doing their best to control territories to drill in their Search (Ng, 2013, p. 8) and to obtain greater digital dividends⁴. This is why Big Data and IoT are often promoted as key tools of market predictions and economic/social dynamics (González & Scherrer, 2015, p. 5). The World Bank has stated the importance of Big Data and open data in development. However, it recognizes that most of the mass data is in private hands (large telecommunications companies and the Internet), which are reluctant to share them for fear, among other things, of jeopardizing their competitiveness (2016, p. 28).

Described as the revolution that will transform how we live, work and think (Mayer-Shönberger & Cukier, 2013, p. 1), the World Economic Forum (WEF) defines Big Data as a collection of data sets so large and complex that they become difficult to process using available database management tools or traditional data processing applications (2013, p. 3). Garriga Domínguez⁵ explains that the concept of Big Data refers to the existence of an enormous amount of data that can be used for different purposes (2016, p. 28).

The IoT has been defined by the European Union (EU) as the next great wave of economic and social innovation enabled by the Internet. This has been valued, by the year 2020, in more than one trillion euros (2016, page 4). The International Telecommunication Union (ITU) defines it as a global infrastructure of the information society, which supports the “burgeoning network of physical objects and devices” that have an Internet Protocol (IP) address, as well as communication that can be activated on the Internet” (2015, p.39). The IoT platform will continuously send huge amounts of data to each node – companies, homes, and vehicles – in real time at any moment. Its operational logic is to optimize the horizontal production between equals, universal access and inclusion, the same qualities that are essential to generate and cultivate social capital in civil society (Rifkin, 2014, p.18). The development of IoT requires not only an Information and Communication Technologies (ICT) infrastructure, but also the appropriate use of large data volumes or Big Data. This is why, with its use for data processing, more is better than less (Mayer-Shönberger & Cukier, 2013, p.58).

ITU recognizes that more than 50% of IoT’s activity is currently focused on manufacturing, transportation, smart cities and consumer applications, but estimates that within five years all industries will have IoT (2015, p. 40), for which large volumes of information will be needed and generated. The majority consists of data that will be generated through our personally identifiable information and that through its treatment will generate income for IoT companies. In addition, it expects that by the year 2020, between 26,000 and 100,000 million devices will be connected as part of the IoT (2015, p.41). While the EU estimates that by the same year there will be 6 billion connections in its territory that will use this technology. For its part, the EMF estimates that by the year 2022, a trillion sensor networks will be embedded in the market (Hult, 2016). All these sensors will contribute to generating and collecting more information about habits, our activities, tastes, feelings, bodies and us will be reflected in data. As a result, the whole volume of information will become increasingly valuable.
3. The Information We Generate

We create the 70% of the digital universe (Garriga Domínguez, 2016, p.29), through our daily interaction with electronic media and services. In this perspective, the digital economy is based on the information that we all generate, although we have not become aware of it.

In the Cisco Report “The Zettabyte Era-Trends and Analysis”, the company recognizes the dramatic growth experienced by Internet traffic over the past two decades. In 1992, global Internet networks carried about 100 GB of traffic per day. Ten years later, in 2002, the sum was 100 gigabytes per second (GBps). In 2012, the sum of global Internet traffic reached 12,000 GBps. Per capita; IP traffic reached 17 GB, up from 7 GB in 2013. While Internet traffic reached 14 GB per capita in 2018, over 5 GB in 2013 (Cisco, 2016).

According to The Internet in Real Time⁷, approximately every 60 seconds, 1354440 GB of data is transferred through the Internet, which generates around 141,780 dollars of profits to the “Internet giants”⁸. Google processes more than 24 petabytes⁹ of data a day, Facebook uploads 10 million photos every hour, and every second 800 million users “upload” more than an hour of video (Mayer-Shönberger & Cukier, 2013, p. 19). Cisco (2016) has estimated that global mobile data traffic grew by 47% by 2015. Reaching 3.7 Exabyte’s¹⁰ per month. In 2015 more than half a trillion (563 million) of mobile devices and connections were added. Growing a total of 7.9 billion, up from 7.3 billion in 2014. In its report Connected World Technology Report (CCWTR), the company estimated that by the year 2015 large data centers would have stored 1.3 zettabytes¹¹ of information (2012).

And while the WB has acknowledged that more than 40% of the world’s population now has Internet access, there are still 4 billion people worldwide who do not have access to this technology. However, the uncontested task of having access to the Internet is one of the goals of the Sustainable Development Goals (ODS) (2016, p. 5). It is therefore estimated that in the coming years, almost the entire world population will be connected. So we are talking about a huge amount of information that will be generated by us with the use of electronic devices connected to the Internet. What will potentiate the scenario that Big Data and IoT poses, i.e. the existence of more information about us.

At present we are constantly transmitting our coordinates, to our electronic devices we deliver our clicks, which allows us to know that we still exist, that we are in constant activity and even measure the variations of our behavior. This is why Big Data is the Holy Grail of behavioral knowledge (Mayer-Shönberger & Cukier, 2013, p. 100). Through data analysis, companies can design personalized advertising strategies for companies. The search for the creation of better services and products is based on the profiling that companies carry out through our habits in the use of devices. For this reason, both the information obtained through social networks and through the browsers can be used to offer publicity based on people and interests in a very effective way (Suárez Sánchez-Ocana, p. 252).

We also contribute to generate profits through advertising campaigns based on user-generated content (UGC), which allow companies to delegate some of the responsibilities of building a brand to their customers. For example, with apps like Trip advisor, Yelp or Amazon, we give reviews and ratings on the quality of services. The contents generated and consumed by users, involve a value, and it is through this that companies improve services to consumers and position themselves with a better category within the market. In a study carried out in 2012, it was determined that 66.3% of consumers in the United States based their purchase decision on the reviews and recommendations of other users (MacKinnon, 2012, p. 18)¹².

---

6 1 gigabyte (GB) is an information storage unit equivalent to 1,000,000,000-byte.
9 1 petabyte (PB) is an information storage unit equivalent to 1015 (1,000,000,000,000,000) bytes.
10 1 exabyte (EB) is a unit of measure of data storage that is equivalent to 1018 bytes.
11 1 zettabyte (ZB) is an information storage unit equivalent to one trillion gigabytes.
12 Amazon proved to be the most popular website to view comments before buying products or services, 66.7% of respondents attribute their review habits to the site. Travel sites were also very popular among participants, 42.1% of participants acknowledging visiting these sites before determining vacation options (MacKinnon, 2012, p.18).
While the Local Consumer Review Survey of 2015 determined that 92% of consumers consult online reviews before making a purchase (Anderson, 2015). For its part, the Interactive Advertising Bureau (IAB, 2016) reported that 75.95% of users generate some response towards the brands that appear in their social networking services. He also concluded that 20.53% of fans/followers suggest or recommend the brand they follow through social networking services. Therefore, says the IAB: “rrss are already the channel of customer service par excellence. They are not a channel of communication, they are a channel of attention” (2016). That is why it is essential that as individuals we are aware of the economics of personal data and their relationship with the options of use and treatment of them. It is necessary to appreciate all of our information, Big Data and IoT to build a frame of reference for its proper use.

The Value of Our Personal Information

In 2014, Shawn Buckles addressed the problem around the monetary value of his information. He decided to sell his data in an auction through the Internet. He received £ 288 from The Next Web page, for his information. Among the information he put on the sale was: personal profile, conversations via email, online conversations, his thoughts, consumer preferences and his web browsing history13. Referring to the causes for the auction, he argued: “I’ve read that a person’s data is worth less than 50 cents at the moment, so I guess I’ve added a lot of value to my data. On the other hand, I have sold my most intimate information. I do not know if there is any amount for it “(Ehrenberg, 2014). And this was so, according to a report in The Financial Times, advertisers were willing to pay $ 0.0005 per person for general information, such as their age, sex and location, or $ 0.50 for the same data of 1,000 people (Steel, Locke, Cadman, & Freese, 2013). In the year 2013 the newspaper published on its website an interactive calculator14, which allows to explore the value of our data.

For the OECD, there is no commonly accepted method for estimating the value of personal data. Some approaches are based on two aspects: (i) market valuations around personal data or other market-related medicines; or (ii) individual perceptions of personal data and privacy (2013, p. 18). Regarding the first, applications created in personal data can provide quantifiable benefits for businesses of around € 330 billion per year by the year 2020. Advertising Revenue Per User (ARPU), which for Google in the first quarter of 2014 was as much as 45 USD, shows the value of personal data, from online platforms, average. Google has achieved a constant ARPU of more than 40 USD on average since the fourth quarter of 2012. For Facebook the ARPU was of 9.45 USD in 2014, 39% more than in 2013. In the year 2014 Facebook obtained revenues for 12,466 million Of USD, of which 92%, or 11,468 million USD corresponded to advertising. While for the first quarter of 2015 were 3.543 million USD (BBC, 2016). For Google, which has always been an important revenue generator, this constitutes more than 90% of its total revenues in the last decade. Considering that in the year 2015 the company had revenues of 21.3 billion USD, of which 19,078 million correspond to advertising (Martínez, 2016).

While for our users, our personal data has the value of a Big Mac or we estimated it in 2013. In a study carried out by Carrascal et al (2013), concerning a measurement on the value of the Personally Identifiable Information (PII) of web-browsing information, they came to this conclusion. By means of price auctions for the data provided during the web browsing of 168 user subjects, who lived in Spain15, carried out for two weeks, information assessments were extracted in different contexts (eight categories16). The results showed that browsing history was valued at around € 7 (approximately USD 10). The age and address between 25 € or 36 USD. When it came to share PII in specific line services, users valued information relating to their financial transactions on social networks, rather than search and purchase activities. The results also showed that users prefer PII goods, including money (32% - 37%) and service improvements (33% - 37%), followed by free services (14% -18%), and directed advertising (3% - 7%). The results of this study showed that, within the most

---

14 Vine. Http://www.ft.com/cms/s/2/927ca86e-d29b-11e2-88ed-00144feab7dc.html#axzz2z2agBB6R (Date of consultation June 20, 2016)
15 Of the 168 study subjects, 94% were Spanish.
16 The categories on which the study was based were: e-mail, entertainment, finance, news, searches, shopping, social and health. The categories used respond to the popular eight that the online ad-networks like Doubleclick used. Among the results were also counted which of these categories were the most and least consulted, with the following results: Searches 82%, Entertainment: 82%, Social: 78%; News: 76%, finance: 75% Purchases: 75%; Email: 64%, Health: 2%. 

---
valued information, the financial and social, on the one of searches and purchases. The study has implications for the monetization of personal information online, since its objective was based on the understanding of the economic aspects of the PII.

Based on the study by Carrascal et al., Staiano et al. (2014) presented the results of a study carried out to determine the monetary value that people assign to different types of PII while being collected by their mobile phones, including Location and communication patterns. In this study we evaluated the PII corresponding to 4 categories: communications, media and apps. Held in 6 weeks, with 60 participants. An ad hoc Java code was developed, which was scheduled to run every night on a secure server in order to automatically generate personalized daily surveys for each participant. Unlike the study by Carrascal et al, the results of the Staiano et al study, resulted in an average of 2 € for the information. The information on location is the most highly valued by users, in some cases the least auctioned (56%), followed by the media (24%), apps (18%) and communications (2%).

More recently, according to a new survey of 5,000 consumers by digital storage company Western Digital, the average consumer values their personal data at £ 3,241 (Curtis, 2016). That report shows that men tend to rate their personal data higher than women, with the average male consumer placing a value of £ 4,174 on their data, compared with £ 3,109 for women. However, women are less willing to “sell” their personal data, with 31% claiming their “priceless” data, compared to 23% of men.

4. Conclusions

Data that are related to or derived from us can not be considered value-neutral. Therefore, we first need to be aware of the simplicity with which our information is collected. Remember that we have become a data driven society. Second, we need to become aware of the data itself and its implications. To leave the world of the tangible, of the physical damages and to see towards the intangible thing and the repercussions that the treatment of our personal information in us has.

We must consider that the voluntary delivery of our information is the price to pay for access and services, but this will not always be the case. There are other possibilities. Today, data technologies allow us to be empowered to be able to collect, store, manage, use and share our information in accordance with our own standards of privacy, comfort and use. This is with the development of user-centred technologies as subject of control and management of their personal data. In this sense, it has been recognized the importance of empowering individuals in a way that allows them to exercise decision making regarding the management of their information (FEM, 2014, p.1). That is why the knowledge and creation of new tools that inform us about the generation and collection of our data, as well as the practices used around them, are essential for the growth of the digital market. As a result, studies and laboratories of Personal Data Stores (PDS) have been promoted. These stores constitute a technology through which individuals sell their personal data to entities interested in

---

17 This was restricted to voice calls made / received, missed calls were discarded. For example, they were informed: “yesterday you made / received 8 phone calls” or “Yesterday, you spoke on the phone with 3 different people”.

18 The individual location referred to a specific place that the participant had visited the previous day. In this way they were informed: “yesterday, at 23:56 you were in Via Degli Orbi 4, Trento”. They were also informed about the total distance traveled in previous days, such as: “yesterday you covered a total of 13 kms away.” In the same way it was considered a question about location: “yesterday you were in 23 different places”.

19 This measurement was made around the making of photographs at a specific time, for example: “Yesterday at 14:23, you took a photograph”, or “yesterday you took 9 pictures”.

20 Running Applications The individual variable includes the date and time and the name of the application running in the foreground. Some of the questions related to this information were: “Yesterday, at 10:23 you were using Firefox Browser”, “Last night, the Google Talk app was used on your device for 82 minutes,” and “yesterday 8 applications were used on your device”.

21 The study was conducted between the weeks of October 28 to December 11 of 2013.

22 Each day at 12 pm participants received an SMS message reminding them of the completion of their survey.

23 Within the study it was determined that users consider the information about their location as the most sensitive. Within the explanations they indicated: “I do not like the idea of being geo-located”, “This type of information is very detailed and very personal”.

---
purchasing them. Its system is based on an effective and explicit informed consent, developed through adequate protection systems, infrastructure and services (Zhang, 2015, p.16). An example of this is the multi-disciplinary project promoted by the Research Council’s UK Digital Economy Program, called The HAT: Hub of All Things. The HAT is the first multifunctional technology platform in the market, installed in houses, that allows to market with personal data by customized products and future services (UW, 2015). This platform is completely controlled by individuals (Ng, 2014, p.7), allowing them to market their contextualized personal information to other stakeholders, thus providing an unprecedented opportunity for companies to understand the context that influences decision making by the consumers. This platform, launched in June 2014, collects information and analyzes the data generated by a group of volunteers, called digital persons zeros (DP0s). Through sensors placed in the objects of their homes, that is, through the IoT, information is collected and integrated with other personal data. The purpose of the project is to discover in depth how we live our lives in relation to the experience with the things and people around us. All data collected by the HAT is owned by the person. By this, the HAT turned into the first digital vault of personal data available as a digital asset for the trade of future personalized products (UW, 2015).

The EU has emphasized that building user confidence is key to economic development, recognizing the need for a more comprehensive and coherent policy on the fundamental right to personal data protection. This has been reflected with the approval of the General Data Protection Regulation (GDPR). This Regulation ensures a single set of applicable rules, which provide a regulatory framework that encourages the development of the PDS, namely: Article 5, regarding the principles regarding the processing of personal data, related to the legality of the established treatment In Art. 6; Article 7, which clarifies the conditions for the validity of consent as the legal basis for data processing, is particularly relevant for a PDS system, where its treatment is most likely to be based on consent; Article 9, which establishes the general prohibition for the treatment of special categories of personal data (sensitive data) and exceptions to this general rule, based on Article 8 of the European Directive 95/46 / EC. Article 17 on the right of cancellation, and Article 18 on the right of portability are particularly relevant for a PDS provider, and for any public entity that tries to facilitate the adoption of a PDS, because a data management model User-centric could enable the full and informed exercise of these rights in an unprecedented way. However, if such rights are to be fully exercised, interoperability between systems is essential.

The era of Big Data and IoT will bring with it new challenges regarding the use and value of our information. As a consequence of this and to promote the digital economy and the growing data ecosystem, it is necessary for users and individuals to understand how we understand and interact with our information. This power should not be left only in the hands of companies and the lex mercatoria. The better we understand how we relate to our environment in different contexts, we can use this information for better decision-making and improve our lives. In addition, we must be able to decide, effectively, what information we generate we want to share and what not. Becoming aware of the fact and consequences of collecting our data, as well as its monetary value, is the first beneficial step toward empowering our information.

5. References

Anderson, M., 2015, 92% of consumers now read online reviews for local businesses!, https://www.brightlocal.com/2015/08/20/92-of-consumers-now-read-online-reviews-for-local-businesses/(Date of consultation: June 22, 2016)

Araujo Carranza, E., 2009, El Derecho a la Información y la protección de datos personales, Porrúa, Mexico.

Bauman, Z., & Lyon, D., 2015, Vigilancia Líquida, Austral, Barcelona.


Buckles, S., For Sale Personal Data, http://www.shawnbuckles.nl/dataforsale/ (Date of consultation: June 20, 2016)
Interactive Advertising Bureau (IAB), 2016, IV Estudio sobre actividad de las marcas en Medios Sociales, IAB Research, Spain.
Martínez, J., Resultados de Google: supera los 21.300 millones de dólares de ingresos y bate las expectativas, El Economista, http://www.elespanol.com/economia/mercados/20160201/98990414_0.html (Date of consultation: June 20, 2016)
Ramonet, I., 2015, El imperio de la vigilancia, Clave Intelectual, Madrid.
Rifkin, J., 2013, La era del acceso. La revolución de la nueva economía, Paidós, Barcelona.
Rifkin, J., 2014, La sociedad del coste marginal cero. El Internet de Todas las Cosas, el procomún colaborativo y el eclipse del capitalismo, Paidós, Barcelona.
Zhang, J., 2015, Final Group Project. Personal Data Stores, Cambridge University.