APRIL 2018

REPORT

Owning my personal data

Design & Production: Laurent Darellis / Atelier 1
Printing: Mailedit
January 2018

OWNING MY DATA

By

Isabelle LANDREAU

Attorney at the Paris Bar,
Doctor of Law in Intellectual Property
and Law of New Technologies

Gérard PELIKS

Cyber-security engineer,
expert in information security
and President of CyberEdu

Nicolas BINCTIN

Professor of Law Schools,
Senior Lecturer at the University of Poitiers

Virginie PEZ-PÉRARD

Teacher-researcher, Lecturer at the
University of Paris II Panthéon-Assas,
specializing in issues of “privacy”
in commercial practices

Under the direction of

Lucas LÉGER

Doctoral student at the CNAM,
Research Director of the
GenerationLibre Think Tank

Statement of responsibility. This report is organized in three parts. Each author has contributed to a specific part, in their specialist field. Their name is indicated at the beginning of each part, thereby distinguishing the work of each of them. GenerationLibre has overseen all these contributions and edited the final text. The authors are not deemed to approve all the parts.

SUMMARY

01

EXECUTIVE SUMMARY

Every day, we accept dozens of cookies on our computers and agree to one-sided terms of use that dispossess us of **our personal data**, including the most intimate ones to a large extent. The GAFA and other platforms obtain their income by **monetizing these aggregated data**, in particular through **advertising.**

However, users get no direct payment for the raw materials they provide, while the value of their personal data is expected to reach 8% of European GDP by 2020. The free of charge nature of the services hides a lawful **plundering** of our data, i.e. of our person.

Data-wise, we still live in the age of serfdom, giving up our production against "free" services. We need to enter into the stage of market economy.

In this paper, the think tank GenerationLibre studies how to establish a system of personal data ownership. Just as the industrial revolution made intellectual property rights necessary, the digital revolution should create an ownership right on personal data. If data is the oil of the 21st century, is it not time to ask who owns the oilfields ?

This legal innovation would change the way the digital ecosystem works, by giving user-producers:

- **The possibility for e-citizens to negociate and conclude contracts** with the platforms (possibly via intermediaries) regarding the use of their personal data, so that they can decide for themselves how to use them.

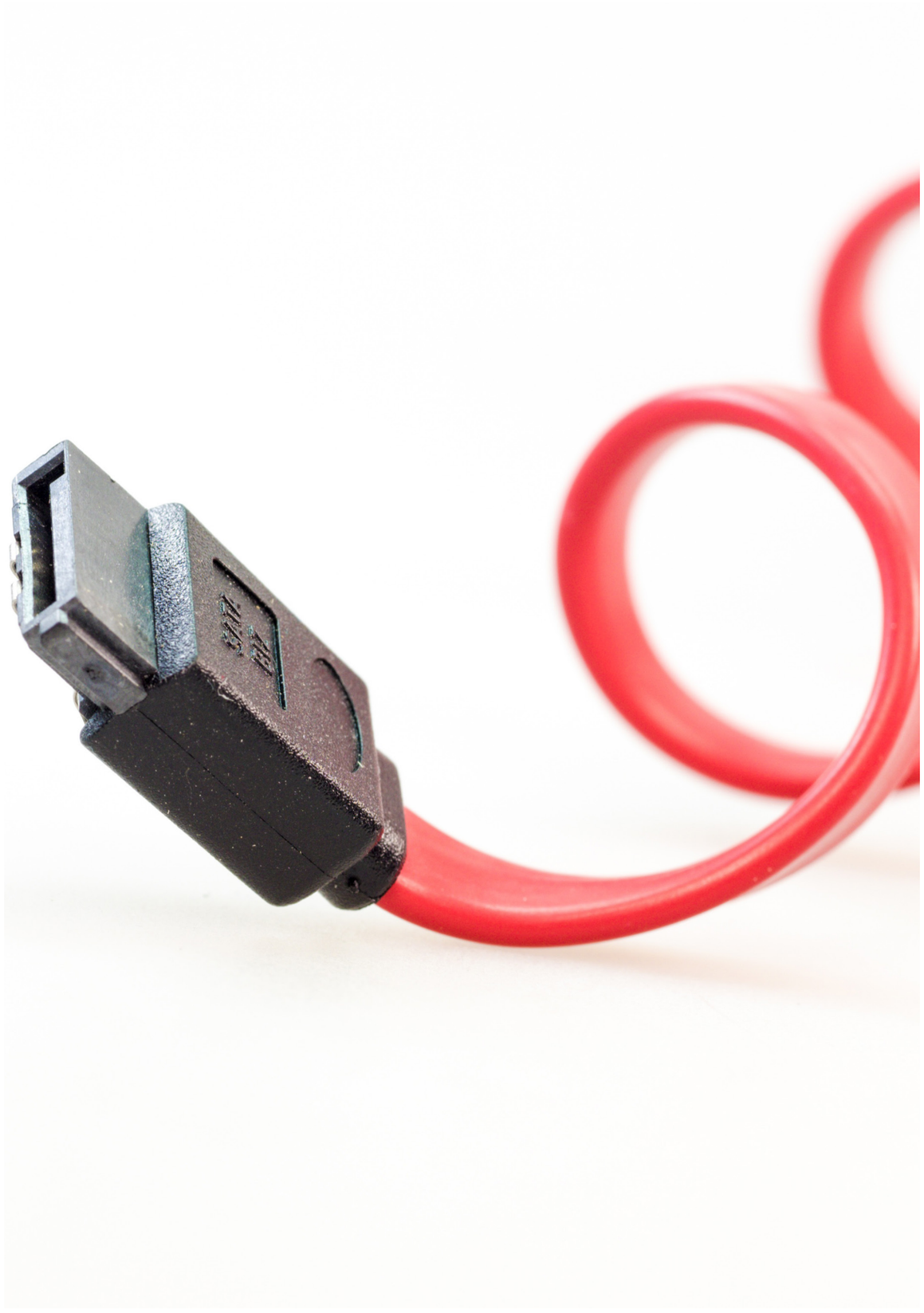
- **The ability to monetise these data (or not)** depending on the terms of the contract (which could include licensing, leasing, etc.).

- **The ability, conversely, to pay the price of the service** provided by the platforms without giving away our data (the price of privacy?).

This paper sets out the principles of the paradigm shift that GenerationLibre advocates, giving back to the individual data producer the legitimate ownership over personal data. This echoes the theory developed by **Jaron Lanier**, author of *Who Owns The Future? (2013)* and co-signatory, together with scholars from Stanford and Columbia universities, of a recent research paper entitled *"Should We Treat Data As Labor? Moving Beyond 'Free'" (2017)*. However, GenerationLibre argues that data should be treated as capital rather than labor, as they effortlessly emanate from self-owned individuals.

Only a proprietary approach will ensure real control on our data by granting the classical rights of *usus*, *abusus* and *fructus*. Only the creation of a data market will be able to rebalance the power relationship between the platforms and their users by providing each one of us with a **real asset**. Although some forms of contractual agreements are already possible in the US, a proper data ownership right does not currently exist in any legal system in the world.

GenerationLibre worked with a team of experts (law professors, engineers, data scientists and economists) to analyze the entire range of socio-economic and ethical issues relating to personal data (I) and to envisage how to introduce ownership of personal data into the law (II).



At the European level, the **General Regulation on the Protection of Personal Data** (GDPR), entering into force on May 25th, 2018, takes a step in the right direction by treating companies as data “guardians” and not owners, as well as by guaranteeing the **portability** of personal data. The establishment of a **property** right would be the logical conclusion of this regulatory progress.

Finally, we explore how technology can now be used to implement this new property right (III). We analyze several possible methods for authenticating users and making their data available. We suggest a blockchain-based model able to manage **“smart contracts”** that would allow everyone to gather and possibly to **sell** their data.

In order to assess the value this new “data market”, our study will be completed by econometric modelling that is being carried out in partnership with researchers at the *Toulouse School of Economics - TSE*. This modelling will provide a better idea of the income that e-citizens could receive from a near-continuous flow of nano-payments.

We believe in a decentralised Internet where **individual identity** is preserved. This is the opportunity for **Europe** to innovate and to propose a **new model in the age of dataism**.

02

INTRODUCTION

The softian bargain

By **GASPARD KOENIG**

Web users are like travellers in past ages, held up by highwaymen. The thieves are the GAFA[1] platforms and other Internet companies.
The larceny: our data.

But unlike our predecessors, we seem to take a certain amount of pleasure in being ripped off. Every day, we allow dozens of *cookies* onto our computers and click “OK” on the “Terms and Conditions” that rob us of our personal data. According to a study conducted by Carnegie Mellon University, the average American signs nearly 1500 of them per year, which would take 76 days to read. PayPal’s terms and conditions are longer than Hamlet. It is simply impossible that under those conditions the web users give their “informed consent”, as the law requires. We cannot read these contracts, let done negotiate them. If we were to scrutinize them a little more closely, we would have good reasons to be wary. A LinkedIn user thus entrusts the social media with the irrevocable right to copy, use and resell all the information it receives. When Facebook undertook, for internal research purposes, to manipulate the emotions of 700,000 users by altering the posts which were displayed to them, the company was able to avail itself of contractual terms allowing it to carry out “research and analysis” on collected data[2]. Mischievously, GameStation had included between the lines of its contracts "the eternal surrender of the user’s soul". Unsurprisingly, 700 users signed up to this Faustian bargain in one day…

Of course, we receive free-of-charge services in exchange for these data. As Jean Tirole emphasizes: *“We often hear that the platforms should pay for the data that we provide them with. In practice, however, some actually do, not in the form of a financial transfer but in the form of free*

1 Common acronym for Google, Amazon, Facebook, Apple.
2 Source: http://www.wired.com/2014/06/everything-you-need-to-know-about-facebooks-manipulative-experiment/.

services."[3] The platforms are in fact paid through the use and/or the resale of these data, primarily to advertising companies that can thereby identify consumers more accurately[4]. Almost all of Facebook's income is thus generated by targeted advertising (increasingly via the application rather than the site)[5], which explains its fierce resistance to ad-blocker software. And this is only the beginning: in a letter dated May 2014 to the *Securities and Exchange Commission*, the American financial regulator, Google explained that advertising would soon appear on fridges, car windscreens, glasses and watches…

The worst dystopias are possible. In a famous scene of the movie *Minority Report*, the hero is offered personalized advertisements while walking in the street (by iris recognition): *"John Anderton! You could use a Guinness right about now"*; *"Get away, John Anderton, forget your troubles."* The deal is quite simple: in return for free access to sites, social media, search engines or pieces of music, we agree to entrust our data to algorithms which offer us in return bespoke products. Let us call this **the Softian bargain** (as in software and Faust). A contractual arrangement between willing parties, but for a mutual benefit? That remains to be seen. The Softian bargain includes at least five one-sided clauses.

A cultural clause: far from being offered an infinite number of possible choices, everyone ends up seeing what they want to see, hearing what they want to hear and reading what they want to read. Based on our past searches, Google delivers first and foremost the information that we want to see, at the risk of increasing our subjective bias[6]. Instead of opening us up to the world, the Internet imprisons us in our bubble[7].

A social clause: we become dependent on "siren servers" which, after having attracted us by the song of simplicity, impose little by little their own system of norms upon us. The banishment of any naked image by Facebook, even if it is a painting by Courbet, is a famous example.

3 TIROLE Jean, *Économie du bien commun*, PUF, 2016.
4 Some, like Criteo, have made a fortune by using "behavioral retargeting" methods.
5 Source: https://investor.fb.com/investor-news/press-release-details/2017/facebook-Reports-Fourth-Quarter-and-Full-Year-2016-Results/default.aspx.
6 It was thus observed, after the Charlie Hebdo attacks, that the dissemination of conspiracy theories was accelerated by Google algorithms (more explicitly, a regular visitor of pro-Palestinian sites will be automatically referred to contents presenting anti-Semitic tendencies).
7 What the online activist Eli Pariser calls the *"filter bubble"*.

As Jaron Lanier, one of the pioneers of virtual reality explains: *"'Free' inevitably means that someone else will be deciding how you live."*[8]

An economic clause: the free aspect is based on the fiction according to which we all have the same value in relation to platforms (namely, zero), even though some users produce large masses of data, and others behave like stowaways, leaving behind only the slightest of digital tracks.

A political clause: If we are currently free to disconnect, there is nothing to ensure that this will be the case tomorrow. The authorities could be tempted to impose *Big Data* on us in the name of the public good. To reduce our electricity consumption, make smart meters compulsory, to prevent traffic jams and accidents, make the interconnection of GPS systems compulsory, to improve public health, make connected bracelets compulsory, etc.

A legal clause, that is fundamental: what kind of privacy am I entitled to in a world where the merest start-up can buy files containing my tastes, my travel, my love-life? It is noteworthy that a champion of the "end of privacy" like Mark Zuckerberg has had a 2-meter high wall built around his Hawaiian residence and is buying up all the houses adjacent to his Palo Alto property. Does the end of privacy only apply to others?

In fact, the Softian bargain is flawed by a basic anomaly: today, we no longer control our data.

However, our data are scattered around the wilderness of the web, owned by nobody, *res nullius*. Someone can appropriate them and convert them into databases that are protected by intellectual property

8 LANIER Jaron, *Who Owns the Future?*, Simon & Schuster, 2013.

rights[9]. This appropriation has a phenomenal value: according to the Boston Consulting Group, the value of personal data in Europe could reach 1,000 billion euros by 2020, or 8% of European GDP. It is the contemporary version of the tragedy of the commons, when the coexistence of several herds on public grazing lands lead to the overexploitation of a scarce resource (grass, in this case). Digital actors graze the same data pasture without worrying about the consequences.

Given this widely accepted observation[10], three kinds of logic can be applied, reflecting three conventional options of political philosophy.

The first is nationalization. The State acquires the pasture and distributes the plots. The data are then considered as a *res communis*, in the same way as air or water in the sea. A kind of national data agency would be necessary to bring together, pool and encrypt all the data of the population. It would then make them available, under certain conditions, to companies which are in the best position to use them. This form of digital communism has quite a lot of support in France[11]. Nevertheless, such a takeover by the State of our data would create a bureaucracy that is diametrically opposed to the culture of the Internet and would give the central power inordinate means of control.

The second option is based on fundamental rights.
The State regulates the use of the pasture land and creates obligations for the shepherds. This so-called "personalist" logic, because it is attached to the rights of the person, has been embraced by the European Commission and the various national regulators (such as the CNIL data protection agency in France). It is based on the concept, defined by the German Federal Constitutional Court of "informational self-determination": everyone should be able to decide autonomously on the use made of their own data. Similarly to the right to be forgotten, the user would be granted additional rights, i.e. the right to allow data to be circulated and to know what use is made of them (including for example via annual reports), the portability of data from one tool to another, the

9 BENABOU Valérie-Laure, ROCHFELD Judith, *À qui profite le clic ? : Le partage de valeur à l'ère numérique*, Odile Jacob, 2015.
10 Conseil d'État, 2014 *Annual Study: Fundamental rights in the Digital Age*, La documentation française, Paris.
11 Particularly under the authorship of Pierre Bellanger, who has put forward the idea of "digital sovereignty".

strengthening of consent procedures, etc. As for the platforms, they would be subject to new obligations: revealing the parameters of the algorithms, offering alternative, non-personalized, services, revealing the processes used to process the information obtained, etc. The risk of this logic is that it would lead to an exponential judiciarization of the digital world, hampering innovation without offering users any real guarantees (and still less payment).

That is why it is urgent to explore a third option, which supplements and in a sense, creates the foundation of the second option: that of assigning ownership to data. In other words, making the people the legal owners of their personal data. That is currently not the case anywhere in the world. The State guarantees the shepherds ownership of their plot of pasture — they are free to exchange them and to find the best balance. If data are, according to the agreed formula, the oil of the 21st century, it is time to ask who owns the oil. In the case of oil it is the primary producer, who resells it to others for refining. Similary, you and me, the data producers, should be paid for the raw material that we supply to the algorithms of *Big Data*. Just as the industrial revolution made intellectual property rights necessary, the digital revolution should create a data ownership right.[12]

This option was explicitly rejected by the French *Conseil d'État* in its 2014 report on digital technology and fundamental rights, on the grounds that it would imply *"relinquishing the logic of protection"* (of the individual by the State). Implicitly, the *Conseil d'État* mentions an argument that is familiar to jurists, namely, that personal data could not be monetized in the name of the protection of fundamental freedoms. Since a person is deemed to be unavailable and cannot be traded, the data which stem from them should also be excluded from the market. They would be part of what *"money can't buy"*, to use the phrase of the American philosopher Michael Sandel.

12 REES Christopher, "Who owns our data?", *Computer Law & Security Review*, 30, 2015.

This moral and philosophical argument calls for at least four points in the rebuttal list :

• The **empirical** argument notes the existence today of huge profits made by data aggregators. Would it not be more fair to distribute the value chain more fairly?

In the name of what sort of “human dignity” should the consumer-citizens be refused their legitimate share of economic production?

• The **legal** argument recalls that ownership rights are designed primarily as a control tool: each person can then dispose of it as they see fit, including rejecting the mechanisms of the market. We can only truly choose to give what we own.

• The **moral** argument justifies abandoning the logic of protection and replacing it with a logic of responsibility: in a mature society, the State must abandon its supervision of the citizen-consumers, and trust them to use their data in an intelligent way.

• The **philosophical** argument asserts the Lockean idea of “self-property”[13] as the ideal of modernity, freeing the individual from the hold of any transcendence.

However, we don't believe that it is necessary to open this sensitive debate. Indeed, data can be considered under the law of property as something that can be appropriated and controlled. In this sense, they remain separate from the person – just as “ideas”, which also are very closely linked to the individual who has produced them, can be intellectual property.

We also need to highlight the concern expressed by the *Conseil d’État* that *“the recognition of an individual’s ownership right to his data creates serious legal issues for the public authorities”* [14] if that was the case,these authorities would have to justify the collection and processing of citizens’

13 LOCKE John, *Second Treatise of Government*, Chapter 5: Property, 1690: *“Though men as a whole own the earth and all inferior creatures, every individual man has a property in his own person ; this is something that nobody else has any right to.“* Locke in fact uses the idea of *“self-property”* already expounded by Richard Overton a few years earlier.

14 Conseil d’État, *Op. cit.*

data for a public interest purpose. For us, this concern represents rather a hope : is it not desirable that the State should explain, if necessary before a judge, how its agencies (including the intelligence services) gather our data? Would this not strengthen the trust needed between the State and its citizens in the digital age?

French and European law is particularly well suited to the inclusion of such an ownership right, the logical next stage in the new regulations regarding data protection. This is the opportunity for Europe to innovate and to impose its model. So that, tomorrow, it will be Facebook that is paying us.

03

PART 1

The socio-economic and ethical aspects of personal data

By VIRGINIE PEZ-PÉRARD, ISABELLE LANDREAU & LUCAS LÉGER

Technological change has often been the source of major economic transformations. The emergence of the IT industry has allowed an effective processing of increasingly larger data sets. The companies that have been able to set themselves up as intermediaries in this new market are now extremely powerful, not only in financial terms, but also in terms of information. For these companies, data has become a commodity and is traded at a premium prices. We think it is essential to ask questions about this new economic order. This introductory section is used to place our analysis within a historical, legal, economic and ethical context.

The goal of this section is firstly and foremost to inform the reader. Consumers can only make choices if they have enough information to be able to measure the impact they have, given the significant amount of time spent on most of the platforms we visit. On average, we spent 2.5 hours per day in 2017 on social media, against 45 minutes in 2012[1]. The longer we stay on these platforms, the more data they gather.

These data-gathering capabilities are today very important. One positive aspect of the Snowden case was to reveal the magnitude of the phenomenon: already in 2010, the PRISM program launched by the NSA could intercept up to 1.7 billion e-mails, telephone calls and other telecommunications[2]. It also showed that the technical capabilities for gathering and processing data are now important.

Once the data are analyzed, internet platforms and governments extract information that goes well beyond simple commercial interaction. We will show the problems this leads to.

1 Source: http://blog.globalwebindex.net/chart-of-the-day/social-media-captures-30-of-online-time/
2 HARCOURT B.E., Exposed, Harvard University Press, 2015.

1. The information society and the data boom

1.1 STARTING WITH THE DIGITAL REVOLUTION

Just like the Industrial Revolution, the Digital Revolution, and is more generally the radical innovations that stem from it, not only changing the way we consume and produce. It is also transforming our entire economic and institutional landscape. Our economies are undergoing a profound change, characterized by the advent of a society of knowledge[3], in which new technologies are helping us to overcome our cognitive limitations. These successive technological revolutions have profoundly affected mankind. *"Technological change is in large part responsible for the improvement in the human condition, whether by the size of its population, the lengthening of life expectancy, the level of education, living standards, changes to work, telecommunications, health care, as well as the effects of human activities on our environment."*[4]

After the Industrial Revolution, we are now witnessing a real break with the past driven by technological change, in which **information has become the value-creating variable**[5]. The networks are the instruments used to transmit and exchange this new form of knowledge, and more widely innovation[6]. Information[7] is transmitted and spreads through the development of software.

3 BELL Daniel, *The coming of post industrial society*, Basic Books, 1973 ; DRUCKER Peter, *Age of discontinuity*, Harper & Row, 1969 ; TOFFLER Alvin, *The third wave*, Bantam Books, 1980.

4 BOSTROM Nick, Technological revolutions: Ethics and policy in the dark, Published in Nanoscale : Issues and Perspectives for the Nano Century, eds. Nigel M. de S. Cameron and M. Ellen Mitchell, John Wiley, 2007, pp. 129–152. Although we are moving away from our subject, it should be noted that there is on consensus on this vision. Indeed, the relationship of mankind to technology was strongly criticized by Jacques Ellul, in *Le bluff technologique*, Hachette, 1988.

5 HIDALGO Cesar, *Why information grows: The evolution of order, from atoms to economies*, Basic Books, 2015.

6 VALENTE, T.W., "Social network thresholds in the diffusion of innovations", *Social Networks*, 18 (1), 1996, pp. 69-89; DEROIAN, F., "Formation of social networks and diffusion of innovations", *Research policy*, 31 (5), 2002, pp. 835-846.

7 Knowledge and know-how are the two fundamental components of the emergence of information.

In a written statement, the investor and co-founder of Napster, Marc Andreessen[8], put forward the argument that **software** was increasingly going to *"conquer the world"* and that all sectors of the economy would be more or less affected by the arrival of new actors able to reduce fixed costs and to provide cheaper and more efficient service. When one takes the time to think about it, the new knowledge economy led to the emergence of software programs to emerge that allow their users to access an efficient service in just a few clicks. We have Amazon in the retail trade, Google in marketing and advertising, drones in the defense sector, the super computer Watson developed by IBM to assist many physicians with their diagnoses, music with Spotify, iTunes and Deezer, payment via PayPal, MOOCs (Massive Open Online Courses) in education, photography, animated films, etc.

Software development is affecting all sectors of the economy and is contributing to automation in industry and services. All these innovations use an ever-increasing quantity of data which most often includes monetizable behaviors. In order to provide such an efficient service, **these new platforms must gather and reprocess personal data** from users who make a purchase on the Internet or simply browse the Web.

1.2 MONITORING BY STATES: "BUSINESS AS USUAL"?

Conventionally, states monitor people, and *Big Data* is the opportunity to monitor them even more, on the pretext of protecting their citizens. **Every second, data flows into "data centers"**[9], most of which are located on American territory. Even though we are fully aware of it, we are to a certain extent **dispossessed of our identity**. The previously obvious dividing line between the private sphere and our public identity (our name, address, etc.) no longer exists. Although Google knows you

8 ANDREESSEN Marc, "Why software is eating the world", *World Street Journal*, August 20, 2011.

9 Source: http://www.iflscience.com/technology/how-much-data-does-the-world-generate-every-minute/. On American territory, 2,657,700 of gigabits of Internet data are generated each minute.

better than your close family[10], the use of mass data is not confined to the GAFA alone. On the contrary, a certain laxity with regard to the ownership of personal data also allows the State to increase its monitoring and to improve its methods. This reversal of the situation has not been brought about coercively.

China is still pioneering in terms of surveillance and restriction of freedom of expression. Already in 2009, it implemented the "Green Dam project". The purpose was to install a parental control software on all new computers produced in China. Although the intention was commendable, a few tests were enough to show that the program went actually well beyond simple control[11]. Besides the security risks that it posed, this software was entirely intrusive and allowed China to carry out intensive espionage of its citizens.

More recently, on June 14th 2014, the Chinese Council of State unveiled its master plan for the construction of a social credit system (2014-2020)[12]. The aim is to build a national reputation system, allowing Chinese citizens to be assessed based on their personal digital data. The goal for the Chinese leadership is to eventually strengthen integrity and *"honesty in public affairs, trade, social issues and the construction of judicial credibility"*[13]. How? By analyzing and compiling the digital behavior of its citizens in the form of individual scores. This control tool, based on a system of incentives to act properly, could currently take the form of a smartphone application that could be deployed in a synchronized version over our various screens, our connected watches, and the part of the *cloud* reserved to us. The establishment of an entire system based on pavlovian reflexes could encourage us to carry out our affairs with complete integrity, rewarded for acting properly and punished for any wrongdoing. Because it is about us and our behaviors that we are talking about here, translated into data shared in the cloud during our

10 Source:http://www.journals.uchicago.edu/doi/abs/10.1086/680084. See also: STEPHENS-DAVIDOWITZ S., *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*, Dey Street Books, 2017.
11 Source: https://opennet.net/chinas-green-dam-the-implications-government-control-encroaching-home-pc.
12 Source: http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion.
13 Source: https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/.

digital travels. It is on the basis of our own data that such systems are established. Without them, they are nothing.

According to a video published by *The Economist*[14], the Chinese Government has developed **a powerful facial recognition software**, particularly through the growing effectiveness of "machine learning" tools. China has thus compiled a database of approximately 700 million different profiles, i.e. half of its population. Companies which manage sensitive content, such as banks or even States could do likewise in the medium to long term. In this way, we could improve the security of individuals. Facial recognition is an effective tool to limit identity theft and hunt down criminals and terrorists. Currently, the start-up Face-six boasts a success rate of 99%. It is still not good enough for security applications, which would potentially return too many false positives to manage administratively, but this rate is growing.

People may argue that this trend only takes place in China or other undemocratic countries. However, the Snowden case revealed that mass surveillance was not a monopoly of governments with questionable democratic principles.

1.3 HUNTING PERSONAL DATA: A BUSINESS MODEL

We are also new seeing commercial facial recognition applications. One of the potential benefits for individuals would be the end of passwords. The iPhone X has deleted them, making access to the device easier. Again, in China, this technology is used for access to amusement parks or as a means of payment in fast food outlets[15]. Behavioural data are thus stored and processed with the aim of improving the "customer experience". Thus, advertising targeted according to our gender and our purchasing characteristics is carried out without us even needing to fill in any form or having a loyalty card.

Does this mean that we control our personal data? Our digital tracks, even if they are not yet used to compile a citizens' social credit score, are

14 Source: https://www.youtube.com/watch?v=nT_PXjLol_8.
15 Again according to the report of The Economist, *ibid.*

definitely used to produce value to those who know how to utilize them. The University of North Carolina estimates[16] the 2012 turnover of the nine largest personal data brokers to be at $426 million. Nonetheless, the primary producers of these data take no part in this market. We share our data **voluntarily** – at least on the surface – most often because it is **required to use some** services on the Internet. Yet, some research is reported to show that we would be willing to pay for our data so it cannot be used by others during our interactions on the Internet[17]. Is this possible? To find out, let us go over why our data are currently at the heart of companies' business models.

In the traditional model, companies fund the digital services they offer to consumers through advertising. That is why most websites or free applications (if not all) include pop-ups or advertising banners that generate advertising revenue. But in the face of resistance of consumers (increase +20% of *adblocks* in 2016[18]), advertising audiences are falling. Better target advertising campaigns become necessary to increase the profitability of the advertising for advertisers. In addition, it is essential to maintain the attention of the targets by limiting the commercial pressure put on them. Companies are betting everything on data. These are collected and used both to develop a detailed profile of the customer, but also to resell them to "partners" (that are actually also customers).

Last names, first names, email addresses, phone numbers, but also products viewed, purchased, interests, habits, etc. All these data are currently one of the major value sources for companies replacing or supplementing reduced advertising revenue.

But what does the customer get in exchange? The gathering of data by companies is above all part of a mercantile logic. Data are a source of considerable income, via the increase in value of advertising spaces sold (thanks to better targeting), and resale to "partners". They are indispensable for ensuring that the digital technologies operate intelligently and relevantly. They allow businesses to customize services, such as keeping the products placed in the shopping cart during a

16 Source: https://onlinemba.unc.edu/blog/data-brokers-infographic/
17 *idem*.
18 Baromètre Adblocks IAB France/ Ipsos – November 2016, http://www.iabfrance.com/content/presentation-de-la-v2-de-letude-ipsos-realisee-pour-liab-france-sur-les-adblocks.

previous visit for example, or recommending products based on the consumer's profile or previous purchases. Data can thus provide users with fluid, simple and practical experiences. Even if thet are not always aware of it, consumers derive **functional benefits** from the use of their personal information (time-saving, comfort, convenience). Sometimes, they may also derive **monetary benefits**. This is notably the case with loyalty cards. By scanning their card when they make their purchases, customers get reductions or benefits, directly or indirectly, in exchange for points that they earn in proportion to how much they spend. Offering these benefits to customers ensures that the customer "will play the game" and will scan their card for each purchase. This is essential to obtain a complete and faithful view of the customer. Such practices are questionable in terms of fairness. Even if the data provide improved services to the citizen-consumers and even if this improvement is a form of value which meets real expectations, **the share of the value returned to customers** in the form of these benefits is probably **relatively low** compared to **the total value that the distributors derive from them** through the resale of the information.

Restoring **the balance** seems essential from an ethical stance, but also to preserve growth and not to reach a point of no return that would result in citizen-consumers being put off from using these services. To make progress on this issue, it is necessary to understand when and how the data are collected, in practice.

1.4 THE COLLECTION AND USE OF PERSONAL DATA IN PRACTICE

Companies take advantage of every opportunity to collect customers' data, and continuously enrich them through multiple approaches.

- **In the physical store, the loyalty card or the "customer file"**

In the physical store, the check-out is now always accompanied by the perennial question from the assistant "Do you have the store's loyalty card?" or "Are you registered in our customer file?". Even if this information is not useful for finalizing the purchase, consumers are

asked for their identity, address or post code, their email address, their telephone contact details, their date of birth, or even the composition of their household or the first names of their children. Each purchase is carefully recorded: products, date, time, store. **Data are then cross-referenced to develop a profile that is as complete as possible of the customer.** If a childcare company wanted to target expectant mothers, it could therefore buy data about consumers who, over the same period, have bought ovulation or pregnancy tests in the store and whose consumption of alcohol has fallen in the past few months. The brand would thus have a good chance of attaining its target using criteria of this kind and at the same time, maximizing its return on investment.

These practices raise questions: besides the income generated by their data unbeknown to them, these methods can encourage consumers into over-consuming and buying products that they do not really need, enticing them with a pseudo-promotion which gives them the illusion of getting a "good deal". In a context where the rate of indebtedness of households continues to grow (+1.1 percentage point between 2015 and 2016 in France[19]), is subjecting consumers and citizens to such temptations reasonable? The issue deserves investigation especially as recent research has shown that certain marketing practices targeting consumers based on their personal data could encourage them to make purchases that they later regret. This can badly affect customers' relationship with the company[20].

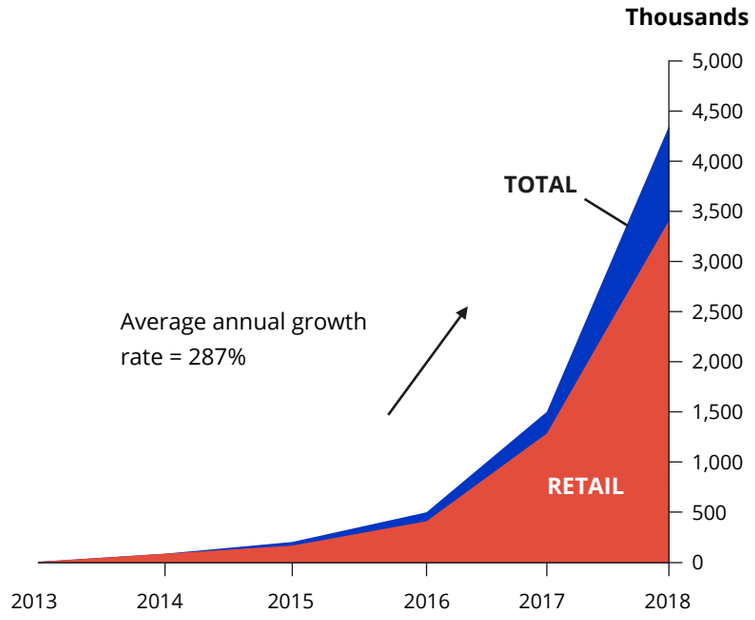
- **In the physical store, *"beacons"* and other connected technologies**

"Beacons" are terminals installed in the stores (in the surrounding area or in the aisles) which use Bluetooth technology. They capture nearby shoppers asthe company is able to link a shopper with their phone, thanks to its customer file. Once the individual is identified near the store or within a specific aisle, **the company can then in real time push them into making purchases with personalized promotional offers**. According to the Business Insider website, in 2018 there will be 4.5 million beacons in service in the United States, 3.5 million for the Retail

19 Figures communicated by the Bank of France, viewable via the following link https://www.banque-france.fr/statistiques/credit/endettement-et-titres/endettement-des-agents-non-financiers.

20 BUTORI R., MIMOUNI A., PEZ V., « Le côté sombre de la pression exercée sur les consommateurs par les programmes de fidélité : enjeux éthiques et pratiques », *Recherche et Applications en Marketing*, Vol. 32, No. 3, 2017, pp. 76-89.

sector (see graph)[21]. 50% of major North American retailers have already launched experimental phases since 2014, and it is estimated that by the end of 2016 85% of their points of sale were equipped. For retailers, this is the fastest technological adaptation since they were equipped with mobile credit card readers! As is often the case, the French market should logically follow the pattern of adoption by the American market by speeding up equipping its points of sale over the next few years.



Forecast of the number of beacons in service in the American market in thousands of units *(Business Intelligence Estimates)*

Physical stores are also capitalizing on digital technologies to provide value-added services to consumers, while at the same time taking advantage of these technologies to gather personal information. In 2017, Facebook tested a new loyalty feature in its mobile application Rewards. This gave users the option of scanning QR codes deployed in physical

21 Source: http://www.businessinsider.fr/us/beacons-are-the-most-important-new-retail-tech-2014-7/.

stores to benefit from targeted promotions[22]. This technology is a useful sales promotion tool for the retailer, which can generate additional footfall in the store. **But it is also for Facebook the opportunity to collect new data about its users and to generate additional advertising income.** This is additionally true for connected loyalty cards, which will in the future offer personalized shopping experiences, based around product recommendations and discounts. That will be the occasion for the retail chains to increasingly expand knowledge of their customers (e.g. the Kiabi clothing chain tested a connected in-store loyalty card in January 2017[23]).

- **The collection of data on the digital customer pathways**

When browsing the web, **each connection is logged**. The pages visited, the pathway followed, the time spent per page, the information read, the products viewed, placed in the shopping cart, and actually purchased, etc. are tracked through *cookies* installed on the devices. Smartphones can track positions by collecting the geographical position of their user several times per minute[24], often via background applications (i.e. without the user being aware of it) fed into the databases of the giants of the Internet.

- **The scanning of private conversations**

Free email services, such as Gmail, are financed through advertising and operate just like dedicated media advertising companies. The purpose of these email services is to **offer targeting as pinpointed and accurate as possible, to allow advertisers to optimize their advertising investments** by maximizing the transformation rate (i.e. the ratio between the number of individuals who have indeed adopted the desired behavior, such as buying the product, compared to the number of ads played/paid for by the advertiser). To offer the most accurate targeting and optimize this rate, the messaging services analyse the private conversations of their web users to extract key words, and thus find out whether they are interested in a particular product, if they intend to go on holiday or make a trip in the near future and where. In short, they identify needs or potential desires.

- **Information gathering in exchange for a free service**

Many sites which are free of charge for the consumer (insurance comparison engines, hotel or plane fare comparison engines, sites for providing property guide prices, used-car guide price sites, etc.) have a business model based on data. To benefit from the service offered by the site, the consumer must provide a number of pieces of information. For flight comparison sites, the user must enter travel dates, destination, number and age of the travellers. For the property guide price site, the user must indicate whether they are a tenant or owner, the condition of the shared parts of the building, or even information which in principle is irrelevant and unconnected with performing the service, such as marital situation and the range of household income. It is easier to understand **why this highly personal information is requested once we know that these sites are largely financed by the resale of this information**. That is why shortly after using these services the user will receive promotional offers for hotels for their next holiday destination or untimely calls from estate agents looking for new customers. Not so free after all …

- **Intelligence watch on social media**

Information may be supplemented by data from social media, in a more or less automated way depending on the relational maturity of the companies. Even if this practice is not widespread today (60% of companies do not incorporate data from social media in their analyses[25]), it is very likely that information will be more widely used in the future as increasing numbers of firms adopt CRM tools (customer relationship management tools) of the latest generation. **This information is very**

22 Source: https://techcrunch.com/2017/05/01/facebook-rewards/.

23 Source: http://www.e-marketing.fr/Thematique/general-1080/Breves/Kiabi-teste-carte-fidelite-connectee-magasin-313408.htm#EABi3g3hu0PSfY8d.97.

24 LEFILLIÂTRE J., "Où étiez-vous hier ? Google peut vous le montrer et c'est effrayant", 19/12/2013 http://www.challenges.fr/high-tech/big-data-comment-la-geolocalisation-de-google-traque-tous-vos-deplacements_10453.

25 Oracle study, *"Can virtual experiences Replace reality?"*, 2016, http://business.lesechos.fr/directions-marketing/marketing/marketing-digital/0211586548032-en-2020-un-marketing-en-realite-virtuelle-303621.php.

valuable: it comes from the individuals themselves who share it believing they are talking with their close circle of family and friends. They are not aware that this information can be bought by brands. This gold mine of information, sometimes very intimate, is still currently little exploited due to the lack of know-how in industrializing its processing. However, its use is expected to intensify in the future.

- **Prize contests**

Prize contests are a simple, fast, and very profitable way of gathering data about individuals by companies. On the pretext of being able to identify you clearly and handing over your prize if you win, taking part in prize contests often requires you to provide a host of personal information. **Sometimes, other information is requested with a promise of offering you a personalized prize, such as the composition of the household or information about your personal preferences, tastes and habits.** Finally you will nearly systematically be given the opportunity to increase your chances of winning by communicating the email addresses of family and friends who may be interested in the offer. By referring them in this way, they themselves will become the target of the contest and provide all this information in turn. Through a "snowball effect", a prize contest campaign can thus pay off in terms of personal information. The information gathered is deemed to be of good quality, because the individuals communicate the information honestly hoping they will be contacted to receive their personalized prize.

- **The purchase of incremental data through customer files**

To better characterize the individuals present in companies' databases, they may buy files to complete with the missing information (or to create new files, as part of canvassing policies). **The files whose quality might not always be guaranteed can be bought very easily from many service providers.** The files are used to increase the value of the information that is already available to the company. Data is valuable when aggregated. By profiling the individual in the most complete manner, the company exponentially increases the value of its data.

- **… and at each contact**

Finally, let us not forget that every contact between the company and individuals is an opportunity to collect data. The letters, phone calls and conversations with sales assistants in the store are scanned, recorded, annotated and "logged" to improve customers' profiles.

Thus, by combining all these sources of information, companies are able to obtain **a very detailed and complete view of individuals**. Although most companies collect a lot of data, there are very few which are currently able to give them meaning. According to an Oracle study[26], 42% of companies are currently unable to extract usable analyses of the data collected. There is thus significant room for improvement in the future. With the new tools and occupations of the *Big Data* era, personalization practices will necessarily intensify.

26 Oracle study, *ibid.*

2. Giving control back to the consumer?

In this picture, citizen-consumers are becoming increasingly worried as they become more familiar with these practices. Some fear the intensification of the “Big Brother” effect and that wealth is being taken away from them. They are beginning to organize the resistance.

What can citizens do in this 2.0 world? Are they recognized as legal persons or coveted objects? It is obvious that the two visions oppose each other: the European vision directed towards the protection of personal data and the American vision directed towards the enormous business potential.

The future of *Big Data* involves not only security but also the establishment of a new business model that places **citizens back at the center of the exploitation of their data**, supported by a simple, practical and existing legal model of which the mechanisms and opportunities we will explain next.

2.1 THE RESISTANCE OF “CITIZEN-CONSUMERS” IS TAKING SHAPE

Consumers have never been so seasoned and aware of business practices. They are familiar with the tools, techniques and business models. At the same time, they are increasingly suspicious vis-a-vis the consumer system and commercial practices. Resistance is gaining ground, taking the form of avoiding certain companies, giving up certain products as well as more active forms such as complaining and boycotts. These movements have always existed but with the advent of digital technology they can be easily organized, quickly visible on a large scale or go viral in just a few clicks. Tools such as the brands’ facebook pages, web forums, consumer opinion sites and even petition sites for coordinated boycott are some of the weapons used to take back power. This is why brands such as Petit Navire, Starbucks, Marineland, H&M or LU, for example, now have to deal with consumers' anger as a real threat.

On a theoretical level, these behaviors result “from a state of motivational resistance”, i.e. an internal tension that consumers feel when they perceive a situation as being oppressive[27]**.** More specifically, these actions are individuals' responses to foil an attempt to pressurize or influence them. In other words, putting too much pressure on consumers could result in obtaining the opposite of what brands want.

Additionally, we are reaching a stage in which consumers are perfectly familiar with the practices of digital economy companies. As such, it creates “marketplace metacognition”[28], preconceived ideas that consumers develop about companies and their tools. That means that they tend to associate their preconceived ideas with all the practices, regardless of their actual characteristics. This knowledge is used to interpret the commercial approaches of companies. When dealing with these companies, consumers seek to maintain their conventional decision-making power, regardless of what the company attempts to obtain through its approach.

Regarding personal data, that is why individuals can consider that their data are appropriated or “stolen” in an unjustified and illegitimate way (for example, to resell them), **even though this use is essential for the functioning of the provided service**. To reduce the risks of such inferences, companies eagerly publish ethical charters to show respect of their customers’ privacy. By providing information about the kind of data collected and their use. They hope to avoid the accusation of dishonesty or selfishness.

Digital citizen-consumers create a real paradox: they want the services offered to be as personalized as possible, tailored to their specific needs and processes to be fluid and “seamless”, without paying a penny. But at the same time, they want relative

27 ROUX D., “La résistance du consommateur : proposition d’un cadre d’analyse”, *Recherche et Applications en Marketing*, 22, 4, 2007, pp. 59-80.

28 WRIGHT P., “Marketplace Metacognition and Social Intelligence”, *Journal of Consumer Research*, 28, 4, 2002, pp. 677-83.

anonymity and absolute control over their personal data. Are these expectations reconcilable?

2.2 TOWARDS PAYMENT FOR PERSONAL DATA?

Given the heightened collective awareness of the practices of companies about personal data and the potential danger that this represents, consumers want to take back control. **88% say they are troubled by the utilization of their personal data;** 88% also say they are worried that their navigation is recorded by private companies[29]. Independent initiatives intended to raise the awareness of citizen-consumers of the advertising income that they generate for the benefit of the Internet giants are flourishing. Such is the Facebook *Data Valuation Tool* [30], a little extension developed by three Spanish researchers that is used to view in real time the total profits generated by Facebook with the user's personal data.

The Facebook Data Valuation Tool extension allows a real-time view of the income generated by Facebook with the users' personal data (http://www.fdvt.org/)

29 Baromètre Adblocks IAB France/ Ipsos – November 2016, http://www.iabfrance.com/content/presentation-de-la-v2-de-letude-ipsos-realisee-pour-liab-france-sur-les-adblocks.

30 Source: http://www.fdvt.org/.

A study published by the Ponemon Institute in 2015[31] shows that **individuals would be willing to share their personal information in exchange for payment**. Their survey, carried out on a panel of consumers, asked respondents how much they would be willing to share specific information. Using this method, the analysts showed that on average individuals valued an piece of personal information at $19.60. The most expensive information is passwords at $75.80, health data at $59.80, payment information at $36, credit situation at $29.20 and consumption habits at $20.60. The least expensive are gender $2.90, name $3.90 and phone number $5.90. Of course the methodology is incomplete since it is only based on individual perceptions, but the report demonstrates citizen-consumers' growing awareness of the value of the assets that they are squandering.

In 2016, **"adblocks"** experienced an unprecedented increase in use (+20%), which shows the desire of consumers to equip themselves with real "shields" intended to block practices impeding their principles or their freedom. The feeling of intrusiveness gets bigger and bigger among citizens but paradoxically they understand that these practices are needed to provide the services they want. How can this paradox be addressed? There are two potential solutions: the first is to **compensate for the psychological and social costs** of data gathering by offering useful or functional benefits; the second is to **pay for the data**.

On the first point, academic research has already looked into the question. It has attempted to identify ways of **reducing the feelings of intrusion** into the private lives of citizen-consumers. Their conclusion is that they are willing to share their personal data if they are given an incentive. Obtaining certain benefits, and particularly utilitarian benefits or ease of use, legitimizes even the most personal use of data. A recent research[32] stresses that companies have two options in this regard: on the one hand, they can offer more benefits to customers to "compensate" for the intrusion, such as an improved interface, greater fluidity, usability, and better personalization. On the other hand, they can better manage the problems of privacy through digital education

31 Source: https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/internet-of-things-connected-life-security/.

32 BELVAUX B., HERAULT S., *"Privacy paradox et adoption de technologies intrusives. Le cas de la géolocalisation mobile"*, Décisions Marketing, 74, 2014.

for both the companies and the users (ethical charters on the use of personal information, trust commitments, education, information and awareness-raising by citizen-consumers in a concern for transparency). On the second point, this report outlins how to come up with **a payment model for citizens** of their personal data.

Future technologies will most probably be even more intrusive. In 2017, specialists forecast increased use of artificial intelligence tools, virtual reality and **"chatbots"**, a kind of software-robot that can join in private conversations to chat with an individual and offer a personalized service. An Oracle study[33] even reveals that more than three-quarters of brands will rely on "chatbots" to manage the customer experience by 2020. These prospects should remind us of the urgency of providing answers to the questions surrounding the value of citizens' data. Given all the services which plan to generate profits by the mass exploitation of citizen-consumers data, it is high time we share out the wealth more fairly and give back to citizen-consumers what belongs to them.

A model rewarding data allows to deal with commercial and social challenges but also raises questions. Rewarding citizen-consumers for their data could result in opportunistic behaviors, where false digital identities (false Facebook profiles, multiple "faxe" email adresses, etc.) could be created for the sole purpose of making profit. This kind of "sabotage" behavior could then devalue the data by making their exploitation less efficient (because based on fake information). In this scenario, the model will probably regulate itself: the data provider is paid based on the effectiveness of the campaigns set up on the basis of their data; it will be adjusted according to the remuneration paid to the citizen-consumer who has transferred his data. The models would take a while to settle down, but quickly regulate themselves, adjust and find a balance. After all, this system pursues a common interest: protecting the business models. Rewarding data in a fair way means preserving growth and lay down the basis of sound, ethical and therefore long-lasting relationships.

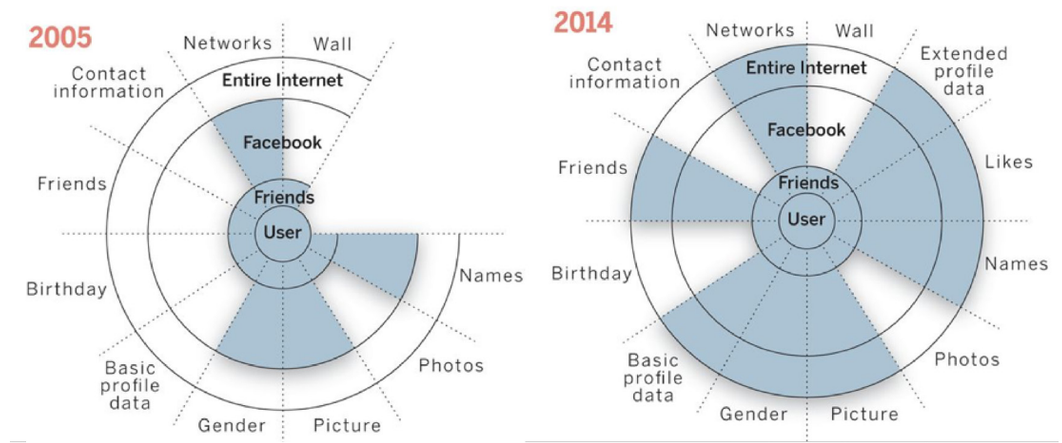
33 Oracle study, *Op. cit.*

2.3 THE RESPECT OF INDIVIDUAL CHOICES, AN ETHICAL CHALLENGE

Personal data are central to governmental and commercial challenges. Practical use of them ranges from the proliferation of military drones[34] to the increasingly accurate targeting of online advertising. For this purpose, Facebook, for example, has continuously increased the visibility of the data available in its default mode, as shown in the next graph.

Besides these aspects, an ethical question arises. What is the place of privacy in the world of *Big Data*[35]? The US General Michael Hayden recalled in an interview that the US government "kills on the basis of meta-data".

Default visibility settings in social media over time



Source: Acquisti, Brandimarte, Loewenstein, "Privacy and human behaviour in the age of information", *Science*, vol. 347 (6221), 2015.

34 Science, *The End of Privacy*, 30 January 2015, p. 497.

35 To better understand the debate from a philosophical point of view, reference may be made to: https://plato.stanford.edu/entries/it-privacy/#ImpInfTecPri.

For its part, the data collected by Google are used to deduce a certain number of things about your individual preferences[36].

This volume of information is in no way comparable, especially as it is sold, with a simple social interaction within a group. It is true that each of us has a public identity on the web. But the GAFA's level of "knowledge" of our behaviors and preferences is disproportionate.

To what extent is privacy a universal concept?
Must privacy be reduced for internal security matters or because the service provided is free?

Beyond the utilitarian vision which would consist of measuring the pros and cons, can we still hope to be able to protect our privacy?

A strong argument is to say that we should simply not use the services of the giants of the internet. Is it the only alternative although the Internet was born with the promise to decentralize and disseminate information?

These ethical issues are important and are also the source of the thinking that has motivated this report to be written, drawing directly from the work of **Jaron Lanier**, a US activist and developer. We believe that a more ethical model is possible, respecting choices and individual freedom.

• Details of the approach used and definitions

Our proposal is to shift responsibility for the data from the platform or the company to its owner. Rights ensue from this responsibility, allowing owners to better protect their personal data, and tip the

36 Stephens-Davidowitz S., *Op. cit.*

economic power back in favor of the consumer of the services, thus facilitating their protection.

Our approach is essentially legal. To what extent can we associate personal data with an ownership right? This question cannot be dissociated from three other aspects: i) being able to characterize personal data within our legal framework, ii) establishing the value of the data and a way to assess it, iii) how to transfer data securely from one actor of the value chain to another.

This seems particularly important since *Big Data* affects all business sectors: industry, health, transport, education, services, local authorities and even the state. What is at stake is the mass storage of data and their operational use as a central part of the new economy.

In France *Big Data* could account for between 3.6 and 7% of GDP. The real problem, as emphasized by Ms. Antoinette Rouvroy, is not so much the *"inappropriate processing of personal data but rather the proliferation and very availability of digital data*[37]*."*

The first challenge is to properly characterize these data, and therefore give them a sufficiently flexible definition to ensure that they can be included in an analytical framework able to cover all their characteristics.

• What kind of data should be considered ?

We give a broad definition of data:

> Primary data are produced by citizens and include their **identity data** (last name, first name, date and place of birth, home, status) and **sensitive data** (sexual orientation, health, religious affiliation, membership of political groups).

37 ROUVROY A., *"Big Data : l'enjeu est moins la donnée personnelle que la disparition de la personne"*, Le Monde/The Conversation, 22 January 2016.

They can be defined as any stored information, recorded on a digital medium, connected to a person and allowing this person to be identified, stemming from the person or from connected objects oobjects or humanoid over which the person has control (author's definition). These data are at the heart of the data economy.

> Generated data are collected by various entities (websites, internet access providers, platforms, e-commerce companies, institutions, associations, NGOS) for lucrative purposes or not. They are collected by trackers and cookies. They include **consumption data** (buying habits) and **financial data** (payment means, status of loans and financing).

> Aggregated data are analyzed according to a precise objective and for mega-data purposes, using generated data. A private citizen cannot perform this task, as private companies and large, sometimes monopolistic groups. It is the speed and calculation capacity of multiple sources of data that create the end value of the data.

The source is always primary data. The citizen has become the free supplier of the 21st century's wealth.

• Data and information: are there any differences nowadays?

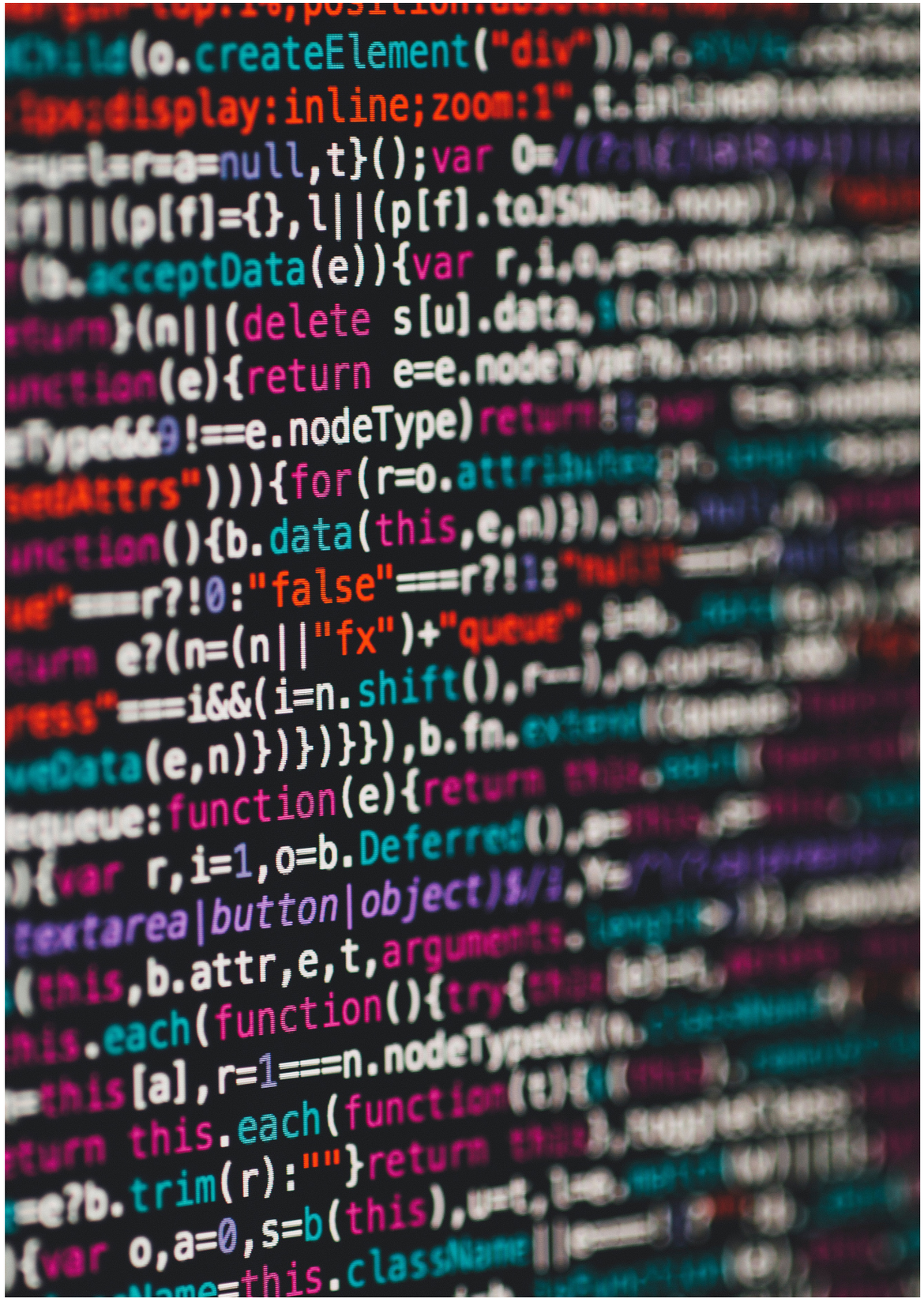
The laws, regulations and European directives since 1978 have included all information and into so-called data mega-data. These data cover a wide range of information such as manufacturing secrets, business information, financial information, patents, trademarks, designs and models, images, writings, lyrics, personal information, family ties, beliefs, consumption habits, etc.

The law distinguishes different regimes for this information. What will characterize the data? Scientifically, a distinction is made between:

- qualitative data
- quantitative data
- categorical data
- countable data

- structured data
- unstructured data

Data are any information that is stored and read by a computer in a digital format, CSV for example. The computer language is not yet as complex as the human language. However, computers are superior in terms of storage capacity and calculation ability. Consequently, aggregate data is where the economic value lies.


(o.createElement("div")),f.
:inline;zoom:1",t.
=null,t}();var 0=
||(p[f]={},l||(p[f].to
.acceptData(e)){var r,i,
}(n||(delete s[u].data,
(e){return e=e.node
!==e.nodeType)
rs"))){for(r=o.attr
(){b.data(this,e,n)}),
r?!0:"false"===r?!
e?(n=(n||"fx")+"queue"
===i&&(i=n.shift(),
(e,n)})})}),b.fn.
queue:function(e){
r,i=1,o=b.Deferr
textarea|button|object)
(this,b.attr,e,t,
is.each(function(){
=this[a],r=1===n.
this.each(
e?b.trim(r):""}
var o,a=0,s=b(this),
Name=this.class

04

PART 2

Creating ownership of data under existing law

Ownership and regulation should not be set against each other[1]. The law provides a legal framework that will legitimize a real market of personal data. This framework will ensure the good faith in the transaction. Here, we review the legal framework in which this new market could operate while ensuring the protection of the consumer.

Our line of argument aligns itself with the GDPR, which guaranteed data portability in its Article 17.

1. Another look at a complex, evolving legal framework

By NICOLAS BINCTIN

The inclusion of data and information in the ordinary or special system of property law has been discussed for thirty years. The original article by Catala[2] which emerged after several initial drafts[3] has since

1 ACQUISTI A., TAYLOR C., and WAGMAN L., "The Economics of Privacy", *Journal of Economic Literature*, 54(2), 442–492, 2016.

2 CATALA P. "Ébauche d'une théorie juridique de l'information", *Rev. de droit prospectif* 1983, No. 1, p. 185; *D.* 1984, chron p. 975; *Le droit à l'épreuve du numérique*, Puf 1998, p. 224 (only this last version is used for the following points); same author, "La propriété de l'information", *Mélanges Raynaud*, Dalloz-Sirey 1985, p. 97.

3 In particular, LECLERCQ P., "Essai sur le statut juridique des informations", Ministry of Justice, 1980; *L'information sans frontière*, (1980), la Doc. Française, Paris ; CHAMOUX J.-P., *Impacts économiques et juridiques de l'informatisation*, Paradoxes 1982, p. 116.

been the source of intense doctrinal debate[4]. It no longer appears necessary to re-examine and discuss these various contributions[5]. Noting the elusive nature of this concept, as well as its central role in the information society, we can simply concur with Catala and remember that *"information is first an expression, a formulation designed to make a message communicable (and) is then communicated",* or may be a *"message of some kind", "intelligible"* and *"communicable"*[6]. The information recorded and communicated takes the form of something[7] legally interesting because of its ability to circulate and to be utilized in a variety of ways[8]. The underlying heterogeneity has no bearing on the analysis that this information may undergo, whether this is sports or stock market results, consumer behavioral data, or of all human populations or otherwise, studied scientifically, medically or sociologically.

The legal approach to data is tricky viewed from the perspective of value theory. Knowing that an anonymous person likes to eat chocolate is a data item but without interest. However, knowing that, in a given population, X people like to eat chocolate is useful information. Data agglomeration creates new significantly interesting information. However, there is no accompanying change in its legal characterization.

Data appropriation mainly comes about through their processing, which begins when they are collected and enhanced, particularly within databases and via algorithmic analyses. The "collection-formulation"[9] concept remains relevant. It has even been clearly laid down through the emergence of the neighboring right of the database producer in the 1995 directive[10]. Freely accessible data are *res communis*, open to everyone, though it does not apply to all data.

The legal status of data means that they can no longer be regarded as facts. They must now be treated within a legal framework. Whether or not they can be appropriated as such, data can be controlled by the party that collects them. Although their collection is free and open, the use of the information collected is subject to the collector's wishes. The status of the data stays the same but the treatment stage changes. The same data can legally be dealt with differently, possibly to the disadvantage of their initial source. For example, when access to data is not public and free but attached to a person. We can distinguish whole data flows (such as road traffic, water or electricity consumption, the frequency of means of public transport), from personal data directly connected to the activity of a given person.

In this second category, collection of the data assumes access to the person and their agreement. The collection is thus made on a contractual basis. It must also be carried out in compliance with the protection of people, and in particular the regulation on the protection of individuals with regard to the processing of personal data and the free movement of these data[11].

After having considered the nature of data (i), we will come back to their appropriation (ii) and the influence of the power of the collector on their status (iii).

4 HILTY R., "La privatisation de l'information par la propriété intellectuelle : problèmes et perspectives. Introduction", *Revue Internationale de droit économique*, 2006/4, p. 353; VIVANT M., "La privatisation de l'information par la propriété intellectuelle", *Revue Internationale de droit économique*, 2006/4, p. 361; same author, "À propos des biens informationnels", *JCP ed. G* 1984, I, No.3132; LUCAS DE LEYSSAC C., "Une information seule est-elle susceptible de vol ou d'une autre atteinte juridique aux biens ?", *D.* 1985, p. 43 ; DEVÈZE J., "Le vol de "biens informatiques", *JCP G* 1985, I, 3210; A. Piédelièvre "Le matériel et l'immatériel. Essai d'une approche de la notion de bien", *Les aspects du droit privé en fin du XX*e *siècle, Mélanges Michel de Juglart, Montchrestien* 1986, p. 55; GEIGER C., "La privatisation de l'information par la propriété intellectuelle. Quels remèdes pour la propriété littéraire et artistique", *Revue Internationale de droit économique*, 2006/4, p. 389; LECLERCQ P., "L'information est-elle un bien ?", *Droit et informatique. L'hermine et la puce*, Masson, 1992 p 91; GALLOUX J.-C., "Ébauche d'une définition juridique de l'information", D., 1994, chr., p. 229; MALLET-PUJOL N., "Appropriation de l'information : l'éternelle chimère", *D. 1997, Chron. 330 ; E. Daragon, "Etude sur le statut de l'information", D. 1998, chron. p. 63;* J. Passa, "La propriété de l'information : un malentendu ?", *Droit & Patrimoine*, March 2001, p. 64.

5 See VIVANT M., "La privatisation de l'information par la propriété intellectuelle", *op. cit.*

6 CATALA P., "Ébauche d'une théorie juridique de l'information", *op. cit.*

7 GALLOUX J.-C., "Ébauche d'une définition juridique.", *op. cit.*; contra, W. Dross, Droit civil – Les choses, LGDJ 2012, No. 483-1.

8 PASSA J., "La propriété de l'information", *op. cit.*, l'information comme "*action* consistant à communiquer à un public des faits ou des opinions".

9 CATALA P., Ébauche d'une théorie juridique de l'information, *op. cit.*, in particular p. 234.

10 Directive 96/9/EC of 11 March 1996 regarding the legal protection of databases, *OJEU No. L 077, 27 March 1996 p. 20*.

11 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

1.1 THE LEGAL STATUS OF DATA

The legal framework governing data can be used to attribute a status to them depending on the nature of the data in question. Two broad categories of datacan be identified based on their nature. Firstly, the long-lasting or short-lived aspect of the data (A); and secondly, the sensitive nature of the data (B), mainly with regard to the situation of the person.

These two categories of data mean we can identify important elements of the regime, both regarding the conditions for collection and the conditions for communication of these data.

A / Short-lived or long-lasting data

The value of the data questions their commodification. Not all data have the same characteristics or undergo the same change in their value over time. This influences their legal statuses. With regard to the temporal consideration of data, certain data only have an economic interest at the time they are issued and potentially collected and disseminated, whereas others may have a much more long-lasting interest.

Short-lived data are characterized by their instantaneity and a necessary control of their disclosure. Collection is controlled and disclosure curbed mainly through the use of systems of responsibility, but also through the emergence of quasi-appropriation. Two sectors illustrate this first category: the betting and gambling sector and financial markets information. In both cases, the data are not appropriated as such but laws regulate the conditions of their collection and dissemination.

There is currently a movement towards commodifying data in order to ensure control over them. Data have no characteristics allowing them to be appropriated but major economic issues relate to them. The tools of ownership are used to establish a legal framework. Data on bets only have a short-lived interest until the outcome of the bet is known. It is only during this time that the law oversees the bets and the dissemination of data about the result.

When the data have no instantaneous pecuniary interest (and a high risk of fraud), collection and dissemination of the data is free, subject to people's rights.

These data are more long-lasting, and their interest is not dependent on instantaneity. Their dissemination of these data will be free and open in equal measure. This is the ideal ground for setting up databases, the value of which is based not on the scarcity of the data but on their collection, accumulation and analysis. The greater the volume of the data, the more the result of their processing is useful. This applies to the data economy, in particular that of search engines and social media. This data category includes collected personal and sometimes anonymized data that allow the creation of behavioral profiles. This is information obtained through loyalty cards or collected through *cookies* on computers or telephones.

The profiles result from the analysis of a host of data, often anonymous, with no individual value in most cases. Their interest lies in finding out the behavior of a social profile which market operators then exploit to tailor their offerings and communications. No specific outcome influences the value of these data, unlike sports betting or stock market information. The reform of Google or Facebook's privacy policies appears to go in this direction[12], offering increased profiling capacity. The information society and its ecosystem mean that people are increasingly aware of the economic importance that these data may represent. Social media are on the front line of this mechanism of collecting anonymous information voluntary-provided.

12 http://www.google.fr/intl/fr/policies/privacy/

B / Sensitive data

The specific status of sensitive data[13] **— referring to the intimacy of a person — is no longer a revolution**. The 1995 EU Directive organized a broad protection and cooperation mechanism within the European Union, before being replaced by Regulation 2016/679. However, since 1995, the technical capacity to collect sensitive data has increased and new categories of sensitive information have emerged, such as tracking by GPS or the geolocation of mobile phones.

The data regime is not guided by its short-lived or long-lasting aspect but by the actual content of the data. The outcome of an operation does not change the characterization of the data. The temporal approach to these sensitive data is a new sticking point, together with the question of the right to be forgotten on the internet.

These personal data are of significant economic sensitivity and economic models are emerging for their use. The case of online social media is significant. Facebook has had to change its practice in the United States and it is likely that it will have to do likewise in Europe. Personal data circulate on social media with uncertainty: firstly, regarding to the perception of the consequence of the action of the person who puts them online; and secondly, regarding the ability of the operator who offers this service to derive a source of income. The economic model of most social media and search engines relies mostly on the provision of a free service in return for the collection of the user's data for commercial use.

It is not a free use of services but an exchange of value: access to searches and media in return for personal data. The economic sensitivity of personal data should mark a change in their legal regime. The 1978 model established by the Data Protection Act, adapted to rapid and transnational circulation in 1995, primarily targeted the control of private data by the states. The second and third changes to this legal framework must allow control on the commercial exploitation of personal data by online non-state operators. The nature of the data remains the same as

13 Other sensitive information, not directly referring to a particular person can be added, in particular information collected during clinical trials to obtain marketing authorisation, which are collected according to the mechanism of Article R 5121-26 of the Public Health Code.

it was over thirty years ago but the risks associated with using them is of a completely different nature.

Away from mistrust of the hegemony and arbitrariness of the state, we must now deal with a mistrust of the online economic operators. This economic sensitivity of personal data does not eliminate the risk to freedom. A judgment of the ECHR of July 2012[14] reminded us of it. The Court clarified the nature of personal data that can be collected during a search ordered by an investigating judge in the premises of a law firm. On the basis of Article 8[15] of the European Convention of Human Rights, the seizure of data must not violate the principle of the seizure being proportional to the purpose of the investigations.

1.2 APPROPRIATION OF PERSONAL DATA UNDER ORDINARY PROPERTY LAW

The value of personal data not only presupposes having means of reservation and appropriation but also means of defense. Besides the protection of personal data by the GDPR, it is necessary to review the solutions provided by property law, to apply them to personal data within a dynamic, ownership approach.

We will consider personal data as an object of property (A) and then the consequences on their utilization and defense (B).

A / Personal data, an object of property

Under the ordinary law of property, data appropriation requires consideration of the concept of "personal data" before invoking the mechanisms allowing this appropriation. It is the possession

14 CECHR, 3 July 2012, *Robathin v. Austria*, application No. 30457/06.

15 Article 8: 1. Everyone has the right to respect for his private and family life, his home and his correspondence. 2. 11 There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

conferred by secrecy or keeping them within the privacy of the person – the control of the data by their initial issuer – which allows one to deduce that an ownership right exists. For the appropriation of property, the relation between ordinary law and special law imposes an appropriate method resulting from the traditional adage, *specialia generalibus derogant* (the specific derogates from the general[16]).

First, it is necessary to check whether the object can be appropriated by special law. If this solution is possible, it rules outs the application of ordinary property law unless the special law offers an option to the owner. Indeed, appropriation made possible by special law applies, but special law may give the owner of the intellectual property the option to fall back on the ordinary property law.

A distinction needs to be made between the possessor who abandons their claim to ownership and a person who seeks to obtain appropriation through the ordinary law. In the first case, the object will not be appropriated but can be freely used by any interested party: it is a *res communes*. For example, for an invention, the special patent law regime allows appropriation of the intellectual property if substantive criteria are met. However, even if these criteria are met, the inventor does not have to use patent law to appropriate their property. Depending on the option available to them, they can also do so using another avenue: secrecy. The invention will then be appropriated under ordinary property law, subject to them retaining the possession under secrecy. Lastly, they can simply disclose their invention before making any patent application and it will become a *res communes*. Conversely, copyright does not offer such an option. If the intellectual property meets the substantive criterion imposed by this ownership regime, it is necessarily appropriated through this means. Exclusion from it is very uncertain.

Where special law does not allow appropriation or if an option is available, and a desire for appropriation exists, the use of the ordinary property law is possible for the appropriation of intellectual property and, more generally, for data, including personal data.

16 See in particular CARBONNIER, *Droit civil – Introduction*, 27th ed. Puf 2002, No. 107, p. 208; CORNU, *Droit civil – Introduction au droit*, 13th ed. Montchestien 2007, No. 329.

Possession must then be exercised over the personal data, an intangible object, effectively enabling ownership to be exercised.

A personal data item can only be possessed if the person who claims to be the possessor holds these two elements: the *corpus* and the *animus*. The *corpus* of possession consists in material acts carried out by the possessor on the object[17]. According to the article 2228 of the Civil Code, these material acts are acts of holding – the object must be subject to the power and control of the possessor – and acts of possession – i.e. the economic use of the object. Secrecy like control enable the object to be held like possession. These two types of acts can be carried out on a data item, which proves the existence of its *corpus*.

Factual power lies in the capacity to maintain it under the seal of secrecy or to control access. The *animus* is independent of the possessed object. This element must be assumed to be the case because it relates only to the behavior of the possessor and there can be no possessory relationship without will[18]. For personal data, the desire to keep them does not require a discussion, so the animus is therefore assumed to be established. *"As regards moveable property, it is through possession that the owner materially asserts his sovereignty to the object by precluding others from taking hold of it"*[19].

Thus, if personal data do not result in appropriation under special property law and if they are subject to possession, then the possessor is able to retain appropriation of the property under the ordinary law, provided they maintain the possessory elements in force. The data are then treated as property appropriated under ordinary law.

Therefore, possession is ownership. This is the last step in allowing the application of the ordinary property law to apply to an intangible object. This allows us to assert that the possession of

17 In this sense, CARBONNIER, *Droit civil – Les biens*, 19th ed. PUF 2000, No. 119 p. 203.

18 In this sense, JHERING, *Études complémentaires de l'esprit du droit romain – Du rôle de la volonté dans la possession*, tome III, 2nd ed. A. Marescq senior, Paris 1891, p. 17 et sq.

19 ZENATI-CASTAING F.and REVET Th., Les biens, *op. cit.*, No. 194

data gives the possessor ownership of these data, in accordance with ordinary property law thanks to secrecy or control[20].

Secrecy, control or intimacy allows possessory hold over intangible property, including data, voluntarily controlled by a person. This possession means that the acquisitive mechanism of Article 2276 of the Civil Code applies and hence the conclusion that personal data are appropriated under ordinary property law.

B / Exercising ordinary law ownership over personal data

Ownership is characterised by its exclusive mechanism. Personal data meet the criteria of this central phenomenon of ordinary law ownership, since possession allowed by secrecy or control is *"a form of expression of the will of the owner to exclude"*[21].

The exclusivity of ownership ensures that its owner can enjoy their property and prevent others from doing so[22]. Exclusivity does not mean that the owner is the only person to have the property in question. Through its behavior, the owner intends to reserve the use of the object. This exclusivity is protected by means of defense that the owner has against infringements of their property. Exclusive power can be used to withdraw all the benefits of the property. *"It is because he enjoys the sovereign power to prohibit his object from others that the beneficiary benefits from all the uses that it may provide, and not by virtue of special prerogatives which are inherent in property law"*[23].

Property law defends exclusivity, in particular by the ability to claim property from a third party. For personal data, case-law has accepted on several occasions the application of a claim for restitution of moveable property against a third party having unlawfully appropriated

20 Application of the mechanism of Article 2276 of the Civil Code to intangible property is the source of a doctrinal debate, but must be allowed. We follow the proposals of William Dross in this area. Since possession is only the factual exercise of a right, it can be applied to intangible property. See DROSS W., *Droit des biens*, 2nd ed. LGDJ 2014, No. 473.

21 ZENATI-CASTAING F. and REVET Th., *op. cit.*, No. 194.

22 DANOS V. F., *Propriété, possession et opposabilité*, Economica 2007.

23 ZENATI-CASTAING and Th. REVET, *op. cit.*, No. 208.

another person's property[24]. By analogy, one must allow the same solutions for personal data. The noting of an ordinary law ownership right results in implications for using and defending the property.

- **Use**

From the application of ordinary property law to personal data, we can draw the following conclusions. All legal provisions apply *mutatis mutandis* with respect to the specificities of this property, in particular the co-ownership regime[25]. Only ordinary law will govern the situation of this property unless the co-owners have intentionally opted for an application of special law to their property. Joint ownership of know-how and secret data is therefore not governed by the patent joint ownership system but by that of ordinary law.

Ordinary law ownership, unlike special ownership for intellectual property, lasts for as long as the possession subsists. It may even be forever, but at the very least it is effective for as long as the secrecy or the control is maintained.

Following the logic of ordinary property law, personal data may be sold, rented or leased like any property for which the owner has a *usus*, a *fructus* and an *abusus*. The courts have also applied other mechanisms of ordinary property law to data whose ownership is appropriated by secrecy, such as claims for the restitution of moveable property[26]. The Court of Appeal of Paris also rightly imposes the application of legal guarantees for peaceful possession when personal data is transferred[27], confirming that the contract includes the transfer of an ownership right, ownership under ordinary and not special law.

In France, the Data Protection Commission (CNIL) is the independent administrative authority responsible for regulating personal data. Currently, the competences of the CNIL for the protection of personal

24 See CA Versailles, 19 May 2006, No. 04/08720, *"the misappropriation of an invention that constitutes one of the cases of the initiation of [claim] proceedings may be successful, with regard to an intangible property claim, even without dispossession of the property, due to an infringement of the enjoyment of the thing over which possession of the property is exercised"*.

25 Cass. com., 7 Dec. 2010, No. 10-30034.

26 CA Versailles, 19 May 2006, Rep. 04/08720, *op. cit.*

27 CA Paris, 11 April 2013, No. 12/21643, see *contra*, J. Passa, *RDC* Dec. 2014, p. 739.

data are broad and its power of sanction is wide-ranging. The maximum amount of the sanctions it can impose have been increased from 150,000 to 3 million euros. These acts concern property belonging to a person who is entitled to exercise prerogatives and imposing choices on the use or circulation of their data.

- **Defense**

Defense is a central aspect of appropriated data. The Business Secrecy EU directive moves in this direction. The bulk of its provisions relate to lawsuits and sanctions, to make sanctions for the infringement of secrecy effective and therefore allow the control of appropriated data. One could also add the solutions of Regulation 2016/679.

Special law has several provisions that can be used to penalize the infringement of control over data.

Still in relation to special law, data are protected depending on where they are stored. If data are included in a database, the law governing the database producer applies to preserve this data. Substantial qualitative or quantitative extraction opens the way to mobilizing the legal arsenal attached to this property regime and obtaining both interlocutory measures and civil or criminal sanctions. Proceedings similar to an action for infringement are available. In the same spirit, if access to a data item results from the violation of an information system, it is possible to act against the perpetrator of this wrongdoing under criminal law[28]. The issue is no longer the data as such but the way in which they have been obtained. Accessing or staying, fraudulently, in all or part of an automated data processing system is punishable by up to two years' imprisonment and a 30,000-euro fine. Thus, criminal law can be applied to data and strengthens the image of a regime for their appropriation through the tools of defense that it provides.

Even more than these special measures, application of the ordinary criminal law is adopted for infringement of property subject to ordinary law. Indeed, although infringement of a data item does not mean that the owner loses possession of it, which precludes them

28 Art 323-1 CP.

initiating proceedings for theft, the owner of a data item can bring proceedings for breach of trust, an accusation relating to property infringements.

1.3 THE POWER OF COLLECTING INFORMATION

The third series of elements that influence the data regime relates to the nature and power of the person who collects the data. Case-law specifically shows that important consideration is given to these elements in assessing the power to collect/disseminate information, especially the ability to benefit from data collected by a third party. In the first place, we can cite the case-law relating to essential infrastructure. In the case of *IMS Health*[29], although a copyright covered the database (which today could also be covered by the right of the database producer), the data collected and used did appear eligible for this appropriation. The solution adopted by the Court of Justice is not very far removed from this classification, which has led several authors to identify a specific category of property or work, informational property, or a data item.

The power of collection is a key element of the legal regime of data. In principle, data that are not secret or not controlled may be freely collected, provided it is possible to have free access to them. Collection is wrongful and should be sanctioned when it infringes a person's privacy and get a sole control over his data. However, some people have a power of collection that enable them to impose communication of data, in general through the exercise of the prerogatives of public power or thanks to de facto monopolies.

The status of collected data can be influenced by the power or the legal nature of the collector (A). This movement has a voluntary correcting system; without calling into question the power of collection and its privilege, the data will then be circulated widely and freely, allowing their reuse by third parties, which will also have an influence on the economic analysis of the data in question (B).

29 ECJ, 29 Apr. 2004, C-418/01, *Rec.* I-5039.

A / The influence of the legal prerogative justifying the collection

The status of the data is influenced by the power of the collector. IThe question is whether the status of the data is different according to whether it is a free collection or the result of the use of the prerogatives of public power. In the latter case, what would be the impact on the status of the data? Consideration is not given to the legal framework allowing or controlling the collection, in accordance with the procedures introduced since 1978, but the status of the data thus collected.

The activity of a public authority is not an economic activity whn it consists of (i) storing in a database data that companies are required to communicate due to their legal obligations, (ii) allowing interested persons to view these data, or (iii) providing them with copies on a paper medium of them. Therefore, this public authority must not consider this activity as an undertaking within the meaning of Article 102 TFEU.

The power arising from the collection and the corresponding capacity to become a database producer and exercise a right over the extraction and reuse of these data does not constitute a change in the analysis of the ECJ. The Court ended it with its decision regarding the ambition of the company *Compass-Datenbank* to obtain from the Austrian State a massive transfer of recent data collected via the commercial register in return for reasonable remuneration, accompanied by the right to reuse these raw data to enable it to offer a service developed with data already accessible to everyone through intermediate agencies. Since the ECJ concluded that the Austrian State was not acting as an undertaking in the market, it did not have to respond to the final preliminary question. The application of the essential facilities doctrine to the case.

Thus, the prohibition on the reuse of data contained in the commercial register falls within the exercise of the prerogatives of public powers and cannot be separated from the other activities of public powers of the state in question. A public authority may legitimately consider that it is necessary, even obligatory given the provisions of its national law, to prohibit the reuse of data contained in a database. The Court appears to refer to the necessary protection, against systematic and organized disclosure.

In these different cases, the legal characterization of the data does not change. Data is not appropriated as such, but the power of the collector or the cause of collection[30] influences their re-utilisation regime.

B / Correcting mechanisms

Correcting mechanisms find their source in the *Open Data* movement. This includes many countries and organizations, including Saudi Arabia, Australia, Austria, Belgium, Canada, South Korea, the US, the EU, Hong Kong, Kenya, Mexico, Morocco, the Netherlands, the United Kingdom, Singapore and Tunisia.

The objective of the public authorities is usually to ensure that **the creativity of developers and entrepreneurs does not come up against legal barriers perceived as obstacles to the development of innovation**. The collective interest prevails over the control by the state of the data that it collects. The state retains the use of a free license applicable to public data put online on the public portal without totally excluding the possibility of paid licenses. With regard to a license, the state affirms that we are considering a right of enjoyment that may legally be granted only to property. Assigning ownership of the data hence appears to be required.

30 See *supra* the status of information included in marketing authorisations.

It is possible to reproduce, copy, publish and transmit the data, disseminate and redistribute, adapt, modify, extract and transform them, in particular to create derived data. It is also possible to use the data commercially, for example by combining them with other data, or by including them in a product or an application.

These huge possibilities depend on the source of the data (at the very least, the name of the "producer") and the date of their last update. The regime here is constructed analogously with copyright and can only strengthen the idea of an ownership approach to data.

2. GDPR: a step in the right direction?

By ISABELLE LANDREAU

Adopted by the European Union on 27 April 2016, the **General Regulation 2016/679 on the protection of personal data (GDPR)** will enter into force on 25 May 2018 and will repeal Directive 95/46/EC on the protection of personal data, in force for the past twenty years. It will also replace the French Data Protection Act.

It has the twofold objective of reaffirming the fundamental principle of the free movement of data – it had become difficult for businesses to determine which national law to apply to their personal data processing – and **to provide a uniform level of protection within the EU** (while avoiding ethical *dumping*, since the level of protection varied depending on the Member State).

Shifting from a declarative system, based in France on the famous "CNIL declarations" that businesses must submit prior to any file processing, **the GDPR pushes personal data rights into the era of *compliance.*** It requires organizations, both private and public, to be able to demonstrate at any time to both the regulator and the people whose data they process that their practices, their processing and their systems comply with a certain number of guiding principles.

In front of each *right* conferred on natural persons, there is an *obligation* for the organization. This paradigm shift is accompanied by a new, far-reaching obligation, since the GDPR requires the establishment of a *"strengthened"* level of security which is *"appropriate to the risk"*. It falls within the protection of privacy and it considers not only the nature, scope, context and purpose of the processing of personal data, but also the risks for the rights and freedoms of persons, the implementation costs and the state of knowledge (Article 32). It means

for the organizations that they have **an obligation to specifically secure personal data** according to the risks involved.

Our project in favor of personal data ownership is in line with the provisions of the Act for a Digital Republic and those of the new GDPR. Indeed, it consecrates at least implicitly the usus and henceforth the *abusus* of personal data, in particular in Articles 26 of the Act and sections 1, 17 and 20 of the Regulation.

Articles 17 and 20 of the GDPR consecrate the fact that personal data are intangible property since the cyber citizen can request to have their data erased (Article 17) and to receive and transmit their data (Article 20: data portability).

The first principle: the free movement of data within the EU.
The circulation of data is neither restricted nor prohibited. In accordance with one of the founding principle of the Common Market, the data becomes a good through its free movement (Article 1). This regulation applies to any establishment located within the EU. Therefore, French subsidiaries of large US groups are subject to the GDPR (Article 3). It applies to natural persons. It is regrettable that this regulation does not consider the connected object of the natural person because it is linked to the use of a citizen and will generate personal data. They must be the property of the citizen user, *"the cyber citizen"* [31], the new man in the digital city (Article 4.1).

Consent must be the linchpin of the categorial utilization of the digital citizen's data. In its Article 7, the GDPR sets out that the responsibility for consent lies with the controller. Currently, **this consent is given by default**. Personal data are siphoned off by the game of acceptance of the general terms and conditions of sale that the digital citizen does not read. It must therefore be a positive act (written declaration Article 7.2). However, any inclination to influence the utilization of your data is wiped out because if you do not agree, the processing remains lawful even if you subsequently withdraw your consent (Article 7.3).

31 Ibid. 11.

Article 17 is interesting because it establishes a "right to be forgotten", which the Court of Justice of the European Union had begun to recognize. The digital citizen may request an end to the dissemination of their data when these data are no longer necessary when the purpose has changed (cf. Article 6-1 a). The withdrawal reason refers to the specific purposes of the data. The system of consent by categorical utilization presented here fits with the scheme of Article 17 1 a) combined with Article 6-1 a). They can also simply oppose this use. In this case, the controller must erase the data without delay. There are also cases of restriction of the use of the data.

Article 20 provides for data portability, i.e. receiving and transmitting personal data to another system (an operator in practice) and the controller may not oppose it. What the arrangements for the transmission and the technical standards are remain to be defined. In practice, data portability is not very effective today. Issues of interoperability of data and processing arise. What about the processing performed by an internet access provider that is useful for my administrative and tax declarations ? Would this information still be accessible if I switch provider ? The operator might lock it technically.

Therefore, portability confirms the *abusus* by the cyber-citizen of their personal data. A confusion must not be made between portability and transfer. Portability does not imply the erasure of the data. The data may still be retained by the operator in line with its processing purpose.

In the GDPR system, there is the obligation to carry out a data protection impact assessment (Articles 33 and 34). There is therefore a case for ensuring utilization of personal data to be able to transform the asymmetric model into a symmetric model of mutual gains in respect of personal data. In addition, the future Data Protection Officer (DPO) simply needs to check that the consent of the cyber-citizen has been expressly given in their missions of Article 39 of the GDPR. The DPO naturally becomes the guarantor of the categorical utilization granted to the cyber-citizen's data.

In addition, the GDPR tightens the obligations between the *"controller"* (the organization on whose behalf the data are processed, and which

sets out the purposes of the processing) and *"sub-contractors"* (the service providers responsible for carrying out the processing according to the instructions of their clients). Up to now, responsibilities could in large part be apportioned contractually between these two key players in processing data, the GDPR requires that controllers ensure that their service providers offer a level of security in line with the requirements seen above and that sub-contractors assist their clients in fulfilling their obligations.

In addition, Articles 73 to 77 provide for actions against the regulatory authorities and the controllers through the traditional channel of legal remedy before the national courts. Legal action by a citizen non-compliant use of the data protection provisions in France has never been taken before a national court until now. It is possible to consider class actions or joint actions through associations.

Finally, the entry into force of the GDPR is accompanied by a considerable increase in the applicable fines. In the event of loss, leak or compromise of personal data, or failure to guarantee the rights set out in the regulation, the sanctions could total **EUR 20 million or 4% of the group's annual turnover**.

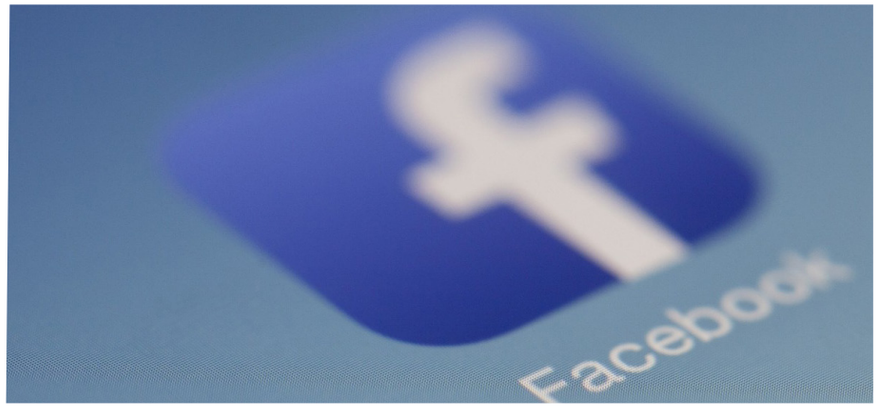
It is true that these new elements do not create a real personal data ownership regime,however they encourage businesses and institutions to establish a genuine "data governance".

Without allowing the citizen to receive payment for their personal data processing, **the GDPR thus shifts the role played by businesses, which are now the *guardians* and no longer the *owners* of these data**. Individuals are again the center of attention, yet while **the value of their information is recognized, the value of their personal data is still out of their hands**.

The question **of aggregated data** remains. The citizens should be able to receive payment on the mass of raw data from those entities which are going to produce aggregated data. Companies in France and abroad are beginning to create *ad hoc* systems to allow payments on data transactions. Several experiments are being carried out in Europe on monetizing personal data, such as MiData in the United Kingdom or MesInfos by the Fondation Internet Nouvelle Génération (Fing) in France.

In the USA, databrokers are already a reality. Databrokers such as ACXIOM[32] or BLUEKAI generate income on the personal data they sell to businesses. Acxiom is reported to have already amassed 600 data items per household from 6 million French households. In principle, Acxiom does not trade sensitive data (e.g. health data). However, the company collects social media data and can create consumer profiles. The Federal Trade Commission (FTC) is thus investigating the data collected by this company

As far as rebalancing the income stream derived from data, it must be recognized that the US system of databrokers is biased in favor of the GAFA. For example, Facebook has signed partnerships with 4 of the biggest databrokers. The current system suffers from a lack of transparency regarding the customers and the way their data are resold. In addition, there are few access rights for US cyber-citizens.



©visualhunt

32 www.zdnet.fr/actualites/data-brokers-aux-etats-unis.

3. Overview of privacy and data protection in the USA

Private data has a different conception in Europe and in the US. The US system focuses on allowing citizens to bring legal action for unfair or deceptive business practices. In Europe, the focus is on privacy protection and control of data.

There are two levels of law requirements in the US: state law and federal law. The former may be more restrictive than the latter.

3.1 PRIVACY AND DATA

A / The concept of privacy

The concept of privacy arose in 1890 in an article from professors Samuel D. Warren and Louis D. Brandeis. It is defined as **a right "to be let alone"**. In 1974, the Federal Privacy Act was passed to regulate government databases and the concept of privacy is recognised as a personal and fundamental right protected by the US Constitution. Since the Clinton Administration, the US Policy towards Privacy Laws is to let the private sector lead the activity. It fits with the US approach of « laissez-faire » in economics as well as in the legal tradition.

The US Constitution does not introduce an explicit right to privacy, on a federal level. However, this right figures in the constitution of many states. That is why there is not a single data privacy act.

There are as many state Privacy Laws as states, and there is not one single data Authority in the US. The Federal Trade Commission (FTC) plays the role of the referral authority for consumers privacy. There are specific privacy laws for what we call sensitive data.

Privacy is also protected by tort law as invasion of privacy, public disclosure of private facts, appropriation or infringement of the

right of publicity or personal likeness and remedies against general misappropriation or negligence.

Adding federal law, states laws and sectoral laws shows that the US regulation covers a wide range of data privacy and use of private data.

B / Classification of data

The FTC considers as personal data any information that can reasonably be used to contact or distinguished an individual. This definition is broader than the European one as it considers any information to identify directly or indirectly an individual.

Sensitive data includes personal health data, credit reports, personal information collected online from children under 13, precise location data and information that can be used for identity theft or fraud.

C / The rules of protection by the Federal laws

The first act is the Privacy Act of 1974, enacted in December 31, 1974. It is called Code of Fair Information Practice and was established to govern the collection, maintenance, use and dissemination of personally identifiable information on individuals when their data is maintained in systems of records by federal agencies. The federal agencies have to keep a system of records in which they collect individual data and each agency must give public notice of their records in the federal register. There is an absolute prohibition of disclosure of information of an individual from this system of records without a written consent of the subject individual.

The Privacy Act states that each agency maintaining a system of records shall :

i) upon an individual's request permit to review the record and to have a copy made of all or any portion in a comprehensible form and

ii) permit the individual to request amendment of its record.

For European citizens, similar rights are stated in articles 15 and 16 of GDPR.

The Privacy Act applies only with the data collected or maintained by agencies. The data collected in records of courts, non-agency government entities are not subject to this Act. On January 25th, 2017, President Trump signed an executive order that eliminates the benefit of the Privacy Act for foreigners for the cause of public safety.

There are many federal privacy laws to protect children. Children Online Privacy Protection Act 1998 (COPPA), financial information with the Fair Credit Reporting Act and the Dodd-Frank legislation, healthcare private information with the Health Insurance Portability and Accountability Act (HIPPA), personal information in the areas of education, cable television, driver's and motor vehicle records, telecommunications customer information, marketing activities, etc.

D / The Example of the Electronic Communications Privacy Act (ECPA), 1986

The Act was passed to promote the privacy of citizens on wiretapping and electronic communication. The ECPA includes the Wiretap Act and the Stores communication Act. It protects citizens against interception of electronic and wire communications, including any oral communication. Anyone who violates the ECPA faces 5 years in prison and fines up to 250.000 USD. Victims are also entitled to file civil law suits to recover actual damages in addition to punitive damages and attorneys' fees.

The fourth Amendment of the Constitution of the United States states that:

"The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."

The Supreme Court considered that individuals have a reasonable expectation of privacy in stored messages (cf Quon v. City of Ontario, CA, 560 U.S. 746, 748 2010). The court also decided that individuals have a legitimate expectation of privacy in their private communications (Nixon

v. Administrator of General Services, 433 U.S. 425, 463- 1977). It was extended to e-mails in 2010 : "a reasonnable expectation of privacy in e-mails stored by their internet providers" (United States v. Warshack, 632 F.3d 266- 6th Cir. 2010).

The judges recognize a right to privacy and in *United States v. Comprehensive Drug Testing,* 621 F.3d 1162, 1180 (9th Cir.), Judge Alex KOZINSKI encouraged data minimization by proposing the following guidelines:

1. Magistrates should insist that the government waive reliance upon the plain view doctrine in digital evidence cases.
2. Segregation and redaction must be either done by specialized personal or an independent third party. If the segregation is to be done by government computer personnel, it must agree in the warrant application that the computer personnel will not disclose to the investigators any information other than what is targeted by the warrant.
3. Warrants and subpoenas must disclose the actual risks of destruction of information as well as prior efforts to seize that information in other judicial forums.
4. The government's search protocol must be designed to uncover only the information for which it has probable cause, and only that information may be examined by the case agents.
5. The government must destroy or, if the recipient may lawfully possess it, return non-responsive data, keeping the issuing magistrate informed about when it has done so and what it has kept.

In 2016, the Email Privacy Act was passed in April to include i) an extension of the warrant requirement to communications stored for more than 180 days and ii) a requirement of notice before email searchers (see www.epic.org)

3.2 OWNERSHIP AND DATA

A / What is ownership under US laws?

Ownership is the right of possession, use, control, and to exclude others of the same, by having a monopole of use on the object possessed. The US recognizes at least three types of property:

• **Real Property.** Real property ownership includes ownership of land, material objects as in feudal times. It is also a right to grant, lease, sell, and exclude others from using its property.

• **Personal Property.** On the contrary of real property, personal property generally is movable or attached to the movable items.

• **Intellectual Property.** It is the property of intangible objects. IP law is statutory and protects inventions (patents), creative works (copyrights), and trademarks. IP law also protects trade secrets, such as the formula for Coca-Cola, or Pierre Hermé's recipe of the cake Ispahan.

Data may be an object protected by trade secrets. Data is all the information stored on an electronic device. Personal Data is any information regarding an individual stored on an electronic format.

B / Who owns the data in the US?

It is easy to say that each citizen owns its data as he generates them. However, with all the channels of communications, platforms, devices, ownership of data becomes more complex.

One option is to divide by channel of communications (Josh MANION, *The power of data ownership: getting it right in 2017*, December 20, 2016). Thus, we will recognise the first party data, data generator: **the citizen** ; then the second party data, data **collectors** and analyzers ; then the third party data, the **marketers**, who will work on cross data and big data.
There is no doubt that data in the US is considered as an intangible asset, as a good in the market. It is treated as a product with a life cycle.

The sharp question is the accuracy of the data.

Another proposal comes from David LOSHIN (*Business-Oriented Data Governance for effective Master data Management* ,2015). He establishes ten possible data owners (quote):

· **Creator** – The party that creates or generates data ;
· **Consumer** – The party that uses the data owns the data ;
· **Compiler** - The entity that selects and compiles information from different information sources ;
· **Enterprise** - All data that enters the enterprise or is created within the enterprise and is completely owned by the enterprise ;
· **Funder** - The user that commissions the data creation claims ownership ;
· **Decoder** - In environments where information is “locked” inside particular encoded formats, the party that can unlock the information becomes an owner of that information ;
· **Packager** - The party that collects information for a particular use and adds value through formatting the information for a particular market or set of consumers ;
· **Reader as owner** - The value of any data that can be read is subsumed by the reader and, therefore, the reader gains value through adding that information to an information repository ;
· **Subject as owner** - The subject of the data claims ownership of that data, mostly in reaction to another party claiming ownership of the same data ;
· **Purchaser/Licenser as Owner** – The individual or organization that buys or licenses data may stake a claim to ownership.

There is a US market place for data and various startups are working on collecting, processing and selling data as Gnip, Infochimps, Windows Azure DataMarket, Factual… They seldom create new data but they clean it up to provide usable data and to connect buyers and suppliers.

The value of the data comes from its potential use and its accessibility. What we call Personal Data Economy (PDE) is emerging in the USA,

where companies can purchase data directly from individuals, like Datacoup.

US citizens have two options: they either pay an extra fee to have their personal data protected (Pay for privacy "PFP" approach) or they accept with no opt-in consent to be paid for the personal data they disclose (PDE approach).

The Federal Communications Commission (FCC) adopted various rules giving the consumer the tools to decide whether or not to share and sell their own data. Nonetheless, these rules have been challenged by the Congress and the Industry trade association.

PDE is close to what the think tank GenerationLibre advocates as each cyber citizen shall have a right to choose whether or not to sell its own personal data. It is what they call the "API of Me" of the "Internet of Me" in the US.

The US approach is a business approach and the question of the ownership of data is now rising as there is no property per se on the personal data.

The business is done by **data brokers** using the digital tools to mine the data, without any form of consent. The main restriction will be the protection of privacy coming afterwards. Data brokers compile data from various sources and sign contracts with other businesses. They have servers to scrutinize consumers data, and without asking any consent they transfer and sell them to third parties. Acxiom (Natasha Singer, Mapping and sharing the customer Genome, NY Times, June 16, 2012) is the biggest one and has 23 000 servers collecting data of millions of citizens. The former Federal Trade Commission (FTC) Commissioner Julie BRILL underlined that consumers lost "control over (their) private and sensitive information".

After the scandal of Cambridge Analytica revealed that 87 millions of facebook users had their profiles disclosed and their data robbed, we should reasonably think of not only the control of our personal data but **also the entire property of our personal data, in order to be able to withdraw, delete, refuse, sell to someone else.**

We are at a turning point facing two options: more citizen privacy or more state security.

President Trump made a choice by signing the so-called « Cloud Act », a law that gives US authorities easier access to data stored abroad. US authorities will be able to demand that internet companies and cloud providers hand over e-mails and other personal information stored beyond the US borders. It will affect Google, Microsoft, Facebook which have data centers in Europe and especially in Paris. It has also an impact of the said privacy shield for data transfer. It completely reduces the need for new negotiations to comply with the European standards of citizen protection and privacy.

To conclude, it is necessary to stick to the guidelines edited by the OECD on personal data in 1980:

1. Notice: data subjects should be given notice when their data is being collected ;
2. Purpose: data should only be used for the purpose stated and not for any other purposes ;
3. Consent: data should not be disclosed without the data subject's consent ;
4. Security: collected data should be kept secure from any potential abuses ;
5. Disclosure: data subjects should be informed as to who is collecting their data ;
6. Access: data subjects should be allowed to access their data and make corrections to any inaccurate data ;
7. Accountability: data subjects should have a method available to them to hold data collectors accountable for following the above principles.

4. Towards ownership of personal data: legal solutions

After having established that the ordinary property system covers data, we offer the concrete proposals which will ensure an ownership right.

4.1 DISTINGUISH DATA FROM INFORMATION

• **Data is an ownership right**

Considering a data item as property means applying property rights to data. According to the law, it must be:

– an object of **desire**, which has a value resulting from its usefulness or rarity;

– an **appropriable** object: the person establishes a relationship of exclusivity with the property, a private situation in relation to third parties;

– an object that can be **transferred**: lawfully passed between individuals, it can be transmitted and is transferable.

Data is currently monetized but has been unilaterally appropriated, excluding the primary provider of the data: the citizen.

Natural and legal persons have assets. We will here attach the data ownership of connected objects and humanoids to the natural person who **uses** the connected object and the humanoid. These will be the only assets of *homo numericus*[33].

33 The phrase of Professor Solange Ghernaouti.

Hence, the citizen has data assets. We will use here Professor Ginossar's broad concept of property[34], which includes the rights and obligations of a legal person in personal assets. It includes animate or inanimate, moveable or immoveable, tangible or **intangible**, present or **future** things from a natural or legal person.

As the provider of the raw material, the citizen is therefore the owner of the primary data and the data generated and aggregated. Those constitutes his personal assets.

He must therefore be able to control them: sell, rent, transfer or even pledge them. It is an enormous opportunity that is being made available to citizens. Let us make not waste it. It would be conceivable to build separate asset funds and to place some of the citizen's data in a trust.

• **A new paradigm: the citizen at the center of the business model for the utilization of data**

The *business model* of the GAFA in which the citizen is a good "sleeping" provider of data with enormous potential can be shaken up. How so? By taking into account the rights of the *homo numericus*[35] and the establishment of a new economic model based on payment of the income generated back to the citizen, the price of their consent to categorical utilization or according to the purpose pursued by their data.

We will therefore switch from a free model to a payment model that will not only be a factor of growth but will also create security. If we stick to the definition of aggregate as given in the 2015 Larousse dictionary, an aggregate is a *"set of elements constituting a whole but not having a defined form, organization, real unity or purpose"*.

34 GINOSSAR S., *Pour une meilleure définition du droit réel et du droit personnel*, RTD civ. 1960, P37
35 Ibid.

Data belong to the person who provides them (classic concept) and the *business model* must be based on the primary data provider: the citizen who will be paid on the added value produced by the data, primary, generated or aggregated.

A / The actors of the data utilization chain

We sketch out the different actors of the data utilization chain with possible business lines:

- **The citizen or the legal person with data assets:** primary data provider, either their personal data, or their entity's data.

- **The data collector** (data centers, ISPS).

- **The data aggregator:** the private entity (commercial or associative) or public entity (possibly a State API, which would collect the categorial data from publicly-owned industrial/commercial establishments) that will have the technical and financial capacity to manage and analyze these data. The citizen must receive income from this collection based on the volume of data and the relevance of the data. The aggregator can sell these data to platforms.

- **The platform:** the status of the platforms must also be reformed. The French Conseil d'État suggests in its annual report a particular status to platforms that *offer their classification or referencing services for content, goods or services put online by third parties* [36].

- **The data retailer**[37] (or analyzer): the retailer or broker is the party that will sell a service linked to the utilization of the data.

- **The Data Privacy Officer** (DPO): the person responsible for personal data in the business, according to the new GDPR. The DPO will become the person responsible for defining the content of usable

36 Conseil d'État, Annual Report 2014,
37 Term used by Gérard Peliks.

data and who could assist with the sale of categorical data according to a strategy defined with the company, in collaboration with the controller.

B / A data utilization business model regulated by the law

Payment by the actors of the data utilization chain may be carried out on the basis of various existing legal mechanisms: in the form of a trademark license agreement, the resale right used by copyright, or even a fee for each utilization granted (pay per loyal use, PPLU[38]) included in the data collection system.

• First model: the trademark and license agreement

According to this mechanism, the citizens register their data assets in the form of a trademark. It is a distinctive sign, a name, a pseudonym, a number (IP address, identity card, health card), which is enclosed within an ownership monopoly.

Utilization of this trademark can only be made under **a trademark license agreement** granted to a third party (e.g. their insurer) and the citizen is paid by a **fee**, on the volume of the data and the use.

The advantage here is that **an upstream ownership monopoly** is created and that the operators must go through the citizen. The downside is that access to the property in this way is not free of charge and creates discrimination between citizens. In addition, the data that are used are not static data. Also, the trademark freezes part of their data at a specific time.

• Second model: the resale right of copyright

The data can be included in **Article L.111-1** of the Intellectual Property Code by considering that it is **a work of the mind** with intellectual, moral and economic attributes.

38 Terminology of the author.

Article 111-2 allows the work to be deemed to have been created independently of any disclosure, by the sole fact of the realization, even incomplete of the author's concept.

Thus, the citizens whose data are absorbed by computer mechanisms such as cookies or others, would be allowed to consider *de facto* that **data stemming from them by their activities on the Internet** are a work of the mind of which content can be utilized on other media.

It will be the responsibility of the data collecting company to inform the web user about the data collected. This is currently often done via an *opt out system*. We should make sure that explicit consent via an *opt in system* would not boil down to a blank cheque by the web user for utilization whose exact content or scope is unknown). The company should also inform the type of utilization of their data (still to be defined), and pay a percentage or a fixed amount on the valuation of the database or of the results of the utilised data to the web user (still to be defined).

Article L 111-3 allows this utilization since intangible ownership is distinct from ownership of the material object. This therefore allows data to be used on many media (TV, laptops, tablets, watches, etc.). It will be sufficient to add to Article L 122-1 of the Intellectual Property Code the concept of "right of collection", as follows: *"The utilization right belonging to the author includes the right of representation and the right of reproduction and the right of digital collection"*.

Article L. 122-3 on the right of reproduction could also be amended by adding a third paragraph: *"For personal data, reproduction is the implicit or explicit collection of data of the web user and his connected objects for a for-profit purpose"*.

Whenever a company utilizes a user's data, it will have to pay a percentage or a fixed amount to the web user producing the raw material.

This might be a controversial measure because it is indeed difficult to find any original feature connected to data or numerical data such as a telephone number, a health card or a credit card[39]. The system here is only attractive for the organization of data usage with, for example, a collective management company of personal data.

• **Third model: the citizen, a creator of a database and the declarative system – declaration of limited use and *pay per loyal use* governed by contract law**

Citizens become database creators because they consent to the human and material investment of which they are the subject-object. **They are the owners of the primary and the generated data**.

An amendment would need to be introduced into the definition of database producer in order to include as the database creator (recognized by the ECJ in its 2015 ruling), the citizen who agrees that his activities generate usable data.

Two conditions need to be met to be recognized as a database producer: It must be a natural person or a company whose registered office/central administration/main establishment is in Europe and who has made a substantial financial, human and material investment (Article L. 341-1 and L. 341-2 of the Intellectual Property Code). The investment here is human. We would therefore change from the database producer to the database creator (citizen).

The declarative system works quite well and has proved itself.
The only drawback is that the mass of data that we will have with connected objects will become unmanageable. The mechanism would be as follows:

39 MATATIA Fabrice and YAÏCHE Morgane, "Etre propriétaire de ses données personnelles: peut-on recourir au régime traditionnel de propriété ?", Revue Lamy de droit immatériel, 2015/114, pp60-63.

1. The database citizen creator declares to the French Data Regulator (CNIL) that he wishes to have his data utilized by a platform data manager and to derive income from them either via connected objects or via platforms or ISPs, for a specific data category and a specific purpose.

2. The database citizen creator consents in writing, in a license agreement to the platform data manager, to the **categorial utilization** of his data according to a **specific purpose**.

This utilization is subject to the consent of the citizen from which the data are derived, for a categorical utilization, limited in time and quantitatively. The declarative system of the data protection regulator (CNIL) is used and makes the citizen the decision-maker of such categorial utilization of the database. This must be valid for any citizen but also open to any company which uses Mega Data[40]. This is a DWYU system standing for *Declare what you use*[41]. This required consent must be explicit. It can be envisaged that the citizen goes onto a CNIL platform (or on a data management API or on a platform of a data managing company platform), where he registers his personal data. This platform would be a center for managing the utilization of personal data. The citizen would receive a categorical utilization request for his data on this platform and would reply to the alert and actively select the checkbox to accept fair and monetized use of its data.

3. The platform manager resells the use of this database to platforms or to ISPs or GAFA through a "smart contact" in the blockchain. Through the blockchain, the purpose and category of data are therefore enclosed and encrypted.

40 INPI study, PI et Économie numérique, 2014.
41 Terminology of the author.

4. The ISP and GAFA platforms pay a percentage to the retailer for the utilization of the data category according to the declared purpose.

5. The data retailer pays an income through micro-payments depending on the value of the use of the data category.

6. This utilization is subject to a payment by the manager of the data platform to the citizen providing the raw material. The payment is made on the platform, which complies with the exceptions of Article L 342-2 of the Intellectual Property Code. The data are stored on the platform which becomes the databroker.

7. A pay per loyal use, or PPLU, is introduced. Control of the utilization of the data is therefore the CNIL's responsibility. Any unlawful and unfair use, about which the citizen would be notified by a system of intelligent alerts, would result in the suspension or withdrawal by the citizen of the selective utilization of their data from the platform or the API.

There would need to be an equal correspondence between the DWYU and PPLU lists and strong oversight by the CNIL, to ensure the respect the rights of the citizen-data creator. The CNIL would remain the authority empowered to impose sanctions in the event of categorical usage that is not in line with the declared purpose.

This system is also easy to set up because it is based on an existing administrative authority, the CNIL, and current mechanisms. It can be included in the data production chain while respecting all the actors.

It becomes part of the new General Data Protection Regulation by the use of express consent and endorses the *usus* and the *abusus* of personal data set out in Articles 17 and 20. In particular, it goes further because it provides for income (*fructus*) for the cyber-citizen.

People may argue that personal data should not be included in assets and must remain non-pecuniary and non-assignable rights, to prevent data utilization misuses and the multiplication of personal data theft[42].

Yet, in the light of the current situation, does excluding data from being considered as assets that can be owned prevent any abusive utilization of a person's data? Are we seeing a fall in the number of identity thefts? Obviously, this is not the case. On the contrary, we notice that cyber-citizens have seen their right of *abusus* taken away from them.

The idea put forward here is to restore an economic balance in an unprofitable game for the cyber-citizen. Furthermore, reducing this asymmetry is one of the aims of the 2016 GDPR. In the proposed system, we see the CNIL as the major control actor to ensure the fair and lawful utilization of data.

42 MATATIA Fabrice and YAÏCHE Morgane, "Être propriétaire de ses données personnelles : peut-on recourir au régime traditionnel de propriété ?", *Revue Lamy de droit immatériel*, 2015/114, p62.

4.2 SHARING THE DATA UTILISATION VALUE CHAIN

By NICOLAS BINCTIN

The data status has an important influence on sharing the value chain. Public data can be freely used and hardly seem to allow the initial data collector or issuer of the data to be included in the value chain. Without excluding an ownership approach to data, the state has ruled out participating in the value chain.

This is not the case with secret data or personal data collected by software solutions when terminals are used. In both cases, the data are not public. They are central to **the transaction**: either they are the direct object of the contract (this is the case regarding a transfer of secret data), or they are granted as one of the conditions of the contract (this is the case of the use of personal data by search engines or social media). The false free-of-charge use of these services is based on the fact that they are based on an exchange of value, content in exchange for personal data instead of monetary consideration. Under these circumstances, we need to find out under what conditions the data issuer will be able to participate in their value chain.

The question is complex and may, in certain ways, suggest that it is possible to **directly monetize data**. For example, for each data item collected, a micro-payment is made for the data issuer and thus a balance is re-established. This idea would be tantamount to saying that the data have value that is greater than the service provided by the online operator because the latter would not only provide the service but also pay the data issuer using the service. Although this situation is not impossible, it would appear to be complex to identify and to implement it. We shall therefore set it aside from the proposals set out below.

At the present stage of development of the online barter of services in return for data, we consider that sharing the value chain must not be viewed from an individual approach but from a collective

approach, more in line with the issue of the mass collection of data. In both cases, the question of data ownership remains an issue because it is a legal tool for the analysis.

By contrast a single data item, mass of data has value in the data economy. Under these conditions, the relevant value is collective and not individual, and part of the value attached to these operations must be collectivized to allow public action to be funded. Therefore, given a barter transaction (services in return for data) between a professional, the search engine or social media, and a consumer, the web user, it is necessary to search for the conditions under which it would be possible to **subject the operation to VAT**. The VAT collected would provide funds to the budgets of the member states of the European Union and would ensure a sharing of the value chain linked to the general interest.

A / A value-added transaction

Barter is an age-old transaction which has found new life in the online economy and the collection of data. Barter between companies, according to the International Reciprocal Trade Association (IRTA), totals between $12 billion and $14 billion per year. It would be necessary today to review this study to include in it the value of trading in the context of the online data economy. Given such exchanges, tax law provides that, if the person is a tax payer, the exchange is regarded as a double sale and so each transfer is subject to VAT. If the exchange is made between two tax payers, the VAT is due only in principle on the profit margin.

Barter is also known to characterize the collaborative economy and local, solidarity currency projects such as the LETS[43]. How these initiatives are considered by the tax authorities is a recurring source of tension, particularly because they are not designed to exclude social contributions and VAT from the payment.

43 MAGNEN J.-Ph. and FOUREL Ch., Study mission about local supplementary curries and local trading systems, *D'AUTRES MONNAIES POUR UNE NOUVELLE PROSPÉRITÉ*, A report delivered to the Secretary of State responsible for trade, artisanal trades, consumption and the social and solidarity economy, 8 April 2015.

In the digital world, exchanges must not be thought in terms of relationships between professionals or individuals, but between professionals and consumer-web users. It is therefore the professional's responsibility to establish the value of the transaction and to declare the appropriate VAT. Indeed, according to the General Tax Code, VAT is due on all *"deliveries of goods and provision of services for an economic activity carried out for a fee"*. This **provision of services** applies to all operations other than supplies of tangible goods, i.e. all intangible services, including the use of a search engine, a social medium, an online films platform, etc.

VAT is due for an economic activity carried out for a fee. It covers market activities, economic operations in return for consideration and exchanges. This consideration is most often represented by a pecuniary element but it can also be in the form of a payment in kind, such as data collection. VAT is thus required to deter companies from preferring bartering to dealing in cash. The intervention on the market must be for a fee, i.e. for fair consideration, which is not solely limited to the search for commercial gain which presupposes, in addition, the search for profit. Even Internet operators which do not have a profit-making objective could still have an action on the market in return for a fee, which could be the case for Firefox for example.

In the light of these elements, there is no doubt that the online collection of data, is a form of barter, which must be described as an economic activity carried out for a fee from a tax point of view.

Therefore, the operators which collect data must declare a value for the transaction and pay the corresponding VAT. If for the web user, the exchange is commutative and reciprocal because it allows access to a search service, and provided that he is aware that the operation is not free but simply provided in return for consideration in kind, it is

imperative that States can collect the value added tax attached to each of these transactions. At a time when the OECD is seeking to implement coordinated tax policies to combat the erosion of the tax base, there is a need to include in the analysis of the value chain, this essential point of the online economy.

To implement such a tax solution, allowing the integration into the local economy of operations conducted online and allowing the states to link their tax revenues with changes to the practices of economic operators, it would be necessary to define a base, a rate and the liable operator.

For such VAT to be applied efficiently, all these points must find a response at the EU level. Not only is VAT subject to European harmonization but the digital economy is also transnational by nature. We believe that VAT should be considered taking into account the mass of personal data collected by the online operators and not simply the data unit collected. We could take into consideration data categories, the presence of cookies, the presence of behavioral analysis tools, etc. For example, geolocation data would have a higher value than those relating to the web user's age.

Furthermore, the implementation of VAT on false free transactions, in the form of bartering services in return for data, would partially offset the effects of the economic models based on this false free service. The cost of the VAT would reduce a little the big margins of certain operators. It would restore more face-to-face and less unfair competition between different online service solutions. It would thus be possible to open up spaces for other economic models for trading and online services and provide a choice for consumers between bartering and purchasing the services they need.

B / VAT Payment in the country of collection

The growth of the digital economy creates challenges related to international taxation. The OECD analyses these challenges in detail[44]. It shows that the digital economy makes it difficult, if not impossible, to ring-fence it from the rest of the economy for tax purposes. It adds however that certain economic models and the essential attributes of the digital economy can exacerbate the risks of erosion of the tax base of the economic operators and describes the expected effects of the measures stemming from all the actions included in the BEPS project. In particular, it presents the rules and enforcement mechanisms that have been defined to facilitate the collection of VAT from the country in which the consumer is located during cross-border transactions between companies and consumers, and which could be used to establish a fair level playing field between domestic and foreign suppliers.

The report explores and analyses possible solutions to the fiscal challenges of a wider scope posed by the digital economy and highlights the need to follow the developments of the digital economy over time. The idea of VAT on collecting and mass processing of data does not appear to conflict with the solutions advocated by this report.

For online commerce, the Commission proposes to improve the VAT environment in the EU, to allow consumers and companies, particularly start-ups and SMEs, to buy and sell goods and services online more easily. Freeing-up the potential of electronic commerce in Europe and creating a digital single market is part of the Juncker Commission's main priorities. The establishment of a pan-European portal for online VAT online payments (the "one-stop shop") should significantly reduce the costs related to compliance with the rules attached to VAT within the EU, estimated at EUR 9,000 per company and per country of declaration, thus enabling companies throughout the EU to save approximately €2.3 billion per year.

The new rules will ensure that VAT is paid in the member state of the consumer, which will result in a fairer distribution of tax revenues between the countries of the EU. These proposals should

44 OECD, *Addressing the Tax Challenges of the Digital Economy, Action 1 - 2015 Final Report*, OECD Publications, Paris.

enable member states to recover the VAT not collected on online sales every year, which is currently estimated at around €5 billion. According to the estimates, revenue losses will reach EUR 7 billion by 2020 if no action is taken. The proposal of imposing VAT on collecting and mass processing of data attached to an online service offer is fully in line with these perspectives.

05

PART 3

Trust technology to rescue your privacy?

By GÉRARD PELIKS & LUCAS LÉGER

As we have explored, users may want to derive financial benefits from their personal data. Tools might seem to be free of charge. However, users are not explicitly informed that their data collected are resold to various organizations such as advertising agencies or companies trying to better target their customers. Users should be able to reject this economic model, even if it means that they will have to start paying the tools or that they woud be paid for supplying their personal data.

It also means that the creators of data have to be able to prove their authenticity and establish that they belong to him when good faith is not sufficient. It is also important that these data are easily accessible to potential buyers and remain integral. Here, we analyze several possible methods for users to authenticate themselves and make their data available, and we propose a model based on a blockchain which manages “smart contracts”.

We will first demonstrate that an **IP address** cannot establish a user’s authenticity for certain, even though it is recognized by the law as personal data. Then we will explain how **an electronic signature** authenticates a user and proves the integrity of a document. Trust is based on the authority which has signed the digital certificate held by this user. Finally, we discuss how **a blockchain** can be used to collect the data of a user. Trust is therefore based on multiple duplication of data and on how many people validate the transactions. Finally, we discuss how **smart contracts** in this blockchain can establish the conditions for the transfer of property between the people who provide their data and the entities which use them, for example to cross-reference the elements collected with other external data, using the algorithms of Big Data.

The model that we propose here assumes significant changes in the tools and methods used and can only be achieved through an

adversarial relationship imposed by the users on the companies developing the tools or exploiting their data.

1. Proving a person's online identity

1.1 THE LIMITS OF THE IP ADDRESS

Is an IP address (Internet Protocol) personal data? The law currently states that it is. But if we analyse this issue at a technical level, we can see that in reality, it is complicated if not impossible to establish a credible correspondence between the IP address of a device and the user who is using it.

Each physical element connected to the Internet has an IP address. This address is a series of four bytes separated by a dot, in the case of IPv4 (e.g. 192.23.34.1) and 16 bytes in IPv6. To make the handling of these addresses simpler to remember, each IP address is associated with a computer name followed by a domain name, for example: node.domain.fr, or even for email: name@domaine.fr. DNS servers (Domain Name Servers) are responsible for establishing the relationship between the representation in IP addresses and the representation in domain names. It is easier for a user to use addresses with domain names which are easy to remember rather than remembering a number. In the text that follows, we characterize an element connected by its IP address alone rather than by the domain names because this add nothing to the relationship between a user and the internet terminal that he uses.

Can an IP address be used to authenticate a user? Not really. It only designates a terminal (PC, tablet, smartphone, etc.) with which the user connects to the Internet. An IP address does not authenticate the user. One might think that if a smartphone belongs to a person, the IP address

of this smartphone is associated with this person. Yet the smartphone may have been borrowed or stolen by another person.

Can an IP address alone be used to ascertain the identity of an internet user behind a terminal ?

In IPv4, the shortage of IP addresses available today means that an individual who is connected is assigned a dynamic IP address for the duration of their session (using the DHCP protocol) by their Internet service provider. This address will then be assigned by this provider to another of its customers. Knowing the time of the connection of a user, the provider can at all times find out to which customer it has assigned a particular address. Indeed, the provider is required to retain at least for a certain amoint of time the connection logs of its customers. **This additional level increases the difficulty of linking a particular person to a particular IP address.** To find out who is behind a particular IP address, in addition to the precise time of the connection, you must also know who the service provider is and whether this provider agrees to make the connection between its customer and the IP address that it has assigned for the duration of their session.

Moreover, in the servers or desktop PCs connected to a company intranet and protected from the Internet by a firewall, an employee will generally access the Internet using the IP address of the outside network controller (the person seeing the Internet) of the firewall. Here, it is their company that assigns, internally, to each server or PC connected to its Intranet, a fixed so-called “RFC1918” IP address, non-routable that begins with a 10 (e.g. 10.20.3.41), by 172 or by 192. The company knows to which computer it has assigned a particular IP address. Outside of the company, on the Internet, there is an added layer which increases the difficulty of attributing a particular transaction to a particular user since all external transactions appear to come from the external IP address of the firewall that protects the organization.

Hence, an IP address cannot identify a person for certain. It can only identify a person’s way of accessing internet, but without having authenticated him. Software layers blur the confirmation of a relationship between a particular person and a particular IP address. **In conclusion, if a person wishes to benefit from its data, they will have to be authenticated by something other than by their IP address.**

1.2 PROVING AUTHENTICITY THROUGH AN ELECTRONIC SIGNATURE

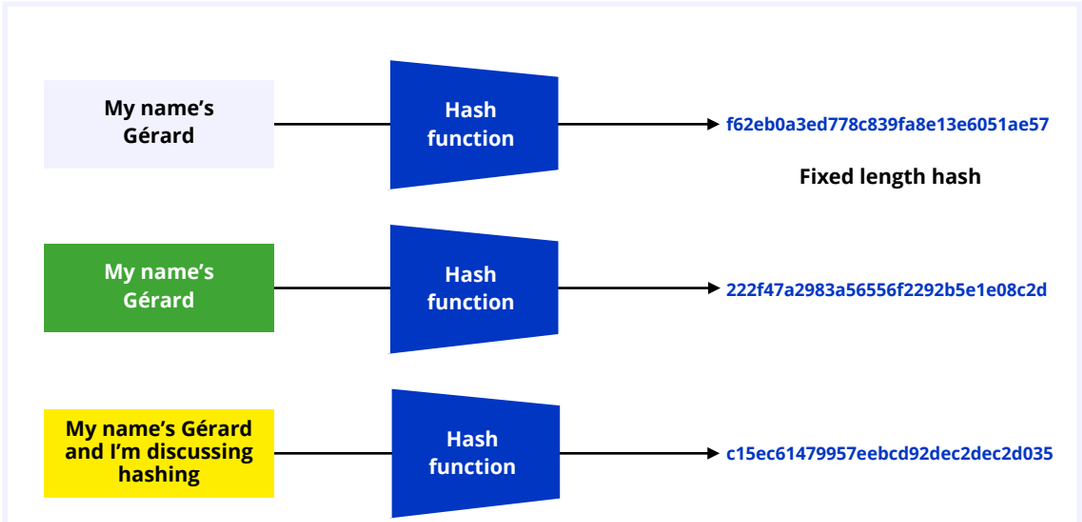
An electronic signature, based on cryptographic mechanisms, can be used to establish a person’s strong authentication regardless of his means of access to digital data, and even if that person is not connected.

We are going to give here some simple explanations about the cryptographic mechanisms used by electronic signatures. We will explain how an electronic signature establishes a relationship between a person and a digital document in a way that is at least as valid, under certain conditions, as a handwritten signature between a person and a paper document can be.

We should understand first what is a hash, and then what is **public key cryptography**, also known as asymmetric cryptography. We will not go into the technical details. We will only give a brief outline, sufficient to understand this vast and complex area. In the next few paragraphs, the main idea is to understand how a document can be electronically signed and how this electronic signature establishes to whom the document belongs.

• **Hashing using hash functions**

A **hash** function, also called a mathematical one-way function, matches a document of variable length with a hash that is a string of characters of fixed length. For example, if this text is put through the hash function SHA256, its hash will be a 256-bit character string, regardless of the length of the text. Let us take another example to illustrate this. If all the works of Emile Zola are processed with the SHA256 function, its hash will also be a 256-bit character string. It would obviously be different from the hash of the text that you are currently reading, but which, if put through this hash function, it would also be a 256-bit character string. If just a comma or any other character were added to one of the texts of Zola, the new hash of the whole of his work will be different from the hash of his initial work.



Does the hash of an original document, attached to this document, establish **the integrity of the document**? Yes, but only for as long as it is not altered and if its hash is subsequently recalculated and attached to this amended document. **Therefore, to establish a document's integrity, this document needs to be linked to its hash. The document is said to be sealed. However, although the integrity of the document begins to be established, it doesn't provide an indication on the author.** To establish this, we will now explain how public key cryptography works.

• **Public key cryptography for encrypting the hash**

Public key cryptography involves two mathematically related keys: a private key and a public key. If we use one of the two keys to encrypt, we can only decrypt using the other. The encryption key and the decryption key are therefore different, and that is why this encryption is called "asymmetric". However, the two keys are mathematically related to each other. Thus, the text you encrypt with one key can be decrypted with the other key. The private key must be kept secret by its owner and he must not disclose it to anyone. By contrast, the public key mathematically linked to the private key, as its name indicates, is public and can be disseminated to everyone. With a public key, it is obviously not possible to reconstitute the corresponding private key. If it were, it would be very costly in terms of calculation and time.

But what establishes that a public key, which can be widely distributed to anyone who asks, really belongs to a particular person who holds the corresponding private key? The answer is that a public key is not provided alone and is included in **a digital certificate** that contains, in addition to the public key, several other elements that identify its owner such as the last name, first name, and possibly an organisation. It also contains the start and end dates of the validity period of this certificate. Above all, it is electronically signed by a trusted authority which proves that the contents of the certificate have not been tampered with. We will see how this is possible once we have explained what the electronic signature is. Remember here that the certificate has been signed electronically by an authority trusted by all these people who will check the electronic signature.

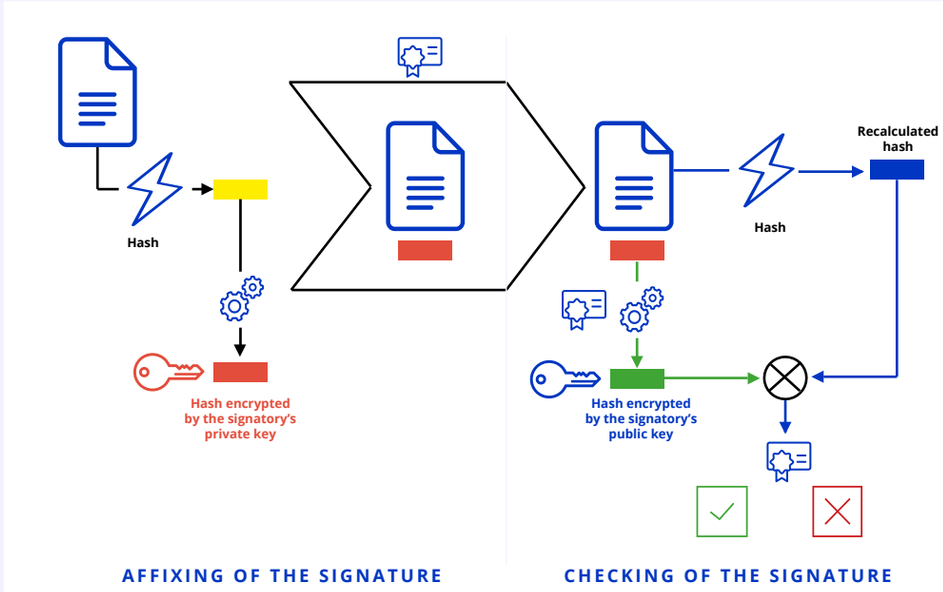
• **The electronic signature to guarantee who a digital data belongs to**

NB: This paragraph is more technical but the reader can refer directly to the diagram below for an illustration of what follows.

The author of a document calculates the hash of their document using **a hash function**. He encrypts this hash using his private key, that only he has, using an asymmetric encryption algorithm such as the RSA or elliptic curves. He attaches this hash to the document. The document is then sealed.

People who wants to establish the authenticity of the owner and at the same time the integrity of the document seek **the certificate containing the owner's public key** by requesting it from him or retrieving it from a directory of certificates. They look at who signed the certificate. If they trust this authority and if the certificate is within the validity dates, they know that the public key that they extract from the certificate is the one mathematically linked to the owner's private key and whose details are found in the certificate. They decrypt the document hash thanks to the public key of the person who has signed the document, using the same asymmetric encryption algorithm. They obtain a decrypted hash. Of course, it is not a user who makes all these calculations, but the application of the electronic signature that he uses, such as an email system or the web browser in the case of a secure HTTPS connection.

The recipient of the signed document then calculates **the document hash** using the same hash function and obtains the hash. If the recalculated hash is the same as the decrypted hash, only the author, whose details have been found in the certificate from which the public key has been extracted and certified by the trusted authority that has signed it, can have encrypted the document hash. This is because they alone have the private key which mathematically matches the public key found in the certificate of the person who has signed the document. The recipient is also assured of the integrity of the document because he is sure that the document has not been changed since its signature, which would have been the case had the recalculated hash not have matched the decrypted hash.



How electronic signatures work

In this way, the authenticity of the author and the integrity of his document are established. The trusted authority has signed the digital certificate of the person who signed the document using the same mechanism as we have described. When a user wants to sign a document, he must therefore have a private key and a certificate containing the public key that matches the private key, since the certificate is signed by a trusted authority.

Nonetheless, not all digital certificates have the same value. It depends on how they have been obtained. Similarly, trusted authorities do not have the same weight, especially in a country which is different from where this trusted authority resides.

An electronic signature is therefore a good way for a user to prove that he is the data owner and that data have not been altered since they were signed, thus proving their integrity. The document has been signed by its owner regardless of the device that he has used. Trust is based on the authority that has signed the digital certificate as it also has a private key and a certificate containing its public key — often included in the certificate store that the browser has.

No-one can therefore modify a certificate signed by a trusted authority without it being noticed, because the result of the hash of this certificate would not match the hash of the decrypted certificate with the public key of the trusted authority.

2. Blockchains to guarantee data authenticity

A blockchain is a "distributed autonomously decentralised database"[1]. The most famous blockchains is the Bitcoin[2], a computer protocol that has resulted in the emergence of distributed applications without trusted third parties, such as cryptocurrencies.

Some see it as a disruptive technology just as revolutionary as the internet and the Web. In the area which concerns us, we will see that **blockchains can provide original solutions** for the economic model that we are discussing here.

Where in cyberspace the data that you want to market should be placed? The simplest thing is to put them in an easily accessible place that everyone can view but that nobody can change. Additional security about the availability of these data can be guaranteed if these data are automatically duplicated on many servers, widely distributed in cyberspace. Blockchain technology provides this possibility, however the limited sizes of the current blocks of the blockchains do not allow large data volumes to be stored. Therefore, we recommend that the blocks are used to store the transferred data hashes only and these will be kept by their owner in a special dedicated portfolio. By using an application such as Zcash, these transfers can even be anonymized.

A blockchain can be viewed as a large registry open to everyone, duplicated, automatically and constantly on many servers with data located in tamper-proof blocks. Indeed, these blocks are chained; each block is time-stamped and depends on the previous ones using cryptographic mechanisms. Once included in the blockchain, a block cannot be changed because of the other blocks in front of them. If one of them is changed, all the blocks that precede it also need to change.

1 Source: https://www.youtube.com/watch?v=wKu_nZcgy3w.
2 Source: https://bitcoin.org/bitcoin.pdf.

Checks on the **integrity of the blocks of the blockchain** will be carried out, on each inclusion of new blocks, by many people called miners. Technically, all the duplications would have to be changed, because they are highly secure against this kind of attack. **The blockchain mechanism can thus be considered as reasonably secure.**

The digital data of the blocks are electrically signed by their owner. If they could be encrypted, they generally are not because everyone must be able to view them. Note also that the owner can identify himself using a pseudonym; if he wants to remain anonymous, only his private key used to include their data in the blockchain will allow him to be authenticated and prove that the data entered are in fact his.

If the digital data of a person who wishes to market them are in a blockchain, the owner of these data can certify that they belong to them until a transaction certifies that the owner has sold them. In that case, the digital data sold no longer belong to them but to the person who has bought them. The blockchain allows everyone to ensure that this is the case.

With the current blockchains standards, the size of each block is limited. The blocks simply contain the hash of the data and their timestamp.

3. Marketing personal data thanks to technology?

- **The self-performing contract to establish the conditions of data sale**

Ethereum, one of the blockchains managing the so-called smart contracts seems to be a solution to manage identities and the digital data that we agree to market.

In the contracts entered in this type of blockchain, which are tamper proof and non-erasable as soon as the miners have validated them, it is possible to define the conditions under which the data can be acquired and how payments must be made. In the case of Ethereum, **payments are made in *ethers***. The *ether* is the unit of one of the many cryptocurrencies found in blockchains. The owners of *ethers* can then convert this cryptocurrency into euros or other currencies via an exchange platform.

In the Ethereum blockchain, the smart contracts are written in a computer language called *Solidity*, that obviously needs to be mastered. It is also essential to predict all cases of application of the contracts because *"code is law"*. However, after the collapse of *TheDao* following a vulnerability in the smart contract code using recursion, *"code is law"* is called into question. Other public or private blockchains will no doubt be created[3] on the Ethereum model, with smart contract description languages that are simpler to use and with better performance, using verification and validation processes by miners that are more suited to our economic model.

3 We can cite *Cardano, EOS, Dfinity or Tezos* as a few potential competitors.

• **Possible use of these tools to manage personal data**

The use of the foregoing implies large changes compared to what is done today. **With this economic model, the data used by the GAFA or by advertising firms will not be directly deducted at source with or without their owner's consent as it is currently done. They will be available on dedicated servers where the owner will make them available subject to certain payment conditions that they will specify in a smart contract.**

The user or the programs they use must supply this blockchain. This is only possible if the process is automated. The tools familiar to us: browsers, office tools will need to be modified to allow users to place their data directly in a blockchain that each user can choose. **If this method were to become universal, the current mechanisms of blockchains will need to be improved to offer the performances required by the avalanche of transactions that will need to be dealt with.**

Of course, we are only discussing here about processing the personal data that a person wants to market. It is better to keep them at home and not to disclose and encrypt data that you want to keep confidential. **It will be better to use tools able to address a blockchain to provide data, rather than using tools that directly provide developers with the data that you want to market.** If your data appear in the Internet ecosystem, although the blockchain managing them makes no reference to this transaction (for example transfer to a third party of your data without your prior agreement), this will prove that your digital data have been acquired illegally. We can then talk about theft or plagiarism.

• **New services for a new profession**

It is obvious that we cannot ask each user to implement this model. We cannot ask them to find the buyer to whom the data can be offered for sale and which are of interest to advertising firms and all organizations eager to process these data to get to know their customers better. **In the chain for making data available, an intermediary is needed between those who produce the data and those who can benefit from them by cross-referencing them with other data and processing them using algorithms to obtain results that they themselves will sell. This intermediary is the "digital data retailer".**

Just like performers use the services of communication managers, users will be able to use the services of digital data retailers who will provide several services. These new service providers will gather data of all types in a *data lake* and be responsible for duplicating and saving the data if they are deemed valuable, with the quality of Veracity and Value in *Big Data*[4].

It is obvious that an isolated data item has little value and that the producer cannot obtain much from it, however a data item cross-referenced with many others may produce information of great value. The cross-referencing of data requiring algorithms and software processing will be made by the digital data retailer.

An additional service will be to render data anonymous or to allow owners to use an alias. Another service of the digital data retailer could be to offer its data providers everything that is necessary for them to sign, in particular the private key and the certificate containing the public key which the user will use to sign their data. The digital data retailer may also be a trusted authority who signs the certificates that it issues.

The retailer also knows where to find customers who will agree to pay a percentage of the sum due to the data producers. The rest is carried out under the terms of the smart contract which will link the producer's obligations and the user of the data. If the idea becomes reality, pushed by users who wish to take advantage of their digital data, a tremendous amount of work will need to be accomplished. These data are often taken and used unbeknown to the users under the pretext that the tool they use (browser, virus protection, etc.) is free.

4 This refers to the famous 5V of Big Data.

When a software is free, it is customary to say that the product is the user. With this model, the user takes back control of the value of its data and can then agree to pay for the software that it previously used for free. This is another business model, and perhaps a model for the future.

• The conditions for success and a few concrete examples

First, our proposal is part of a much broader debate on the use of different technologies to ensure better protection of the privacy of individuals when they go on the Internet. We can distinguish two approaches, which are not exclusive[5]. The first solution is to use technical means to limit the dissemination of personal data. The second tends to set up rights directly related to the use and/or the dissemination of these same data. This short presentation of our solution is based on these two aspects.

Secondly, we note that solutions for monetizing the content of an online site already exist. The concern for protecting privacy or fair reward for producers of content are central to these solutions. With its *Smart Media Token*, the start-up *Steem* rewards and encourages the creation of web content using its blockchain[6].

5 We refer the reader directly to the work of: LE METAYER D., "Whom to Trust? Using Technology to Enforce Privacy", Chapter 17, *in* Wright D., De Hert P. 2016. *Enforcing Privacy Regulatory, Legal and Technological Approaches. Law, Governance and Technology Series*, Volume 25.

6 https://smt.steem.io/smt-whitepaper.pdf. On a fairly similar principle, see also the Akasha project on Ethereum: https://akasha.world/. This is concerned with protecting users' privacy via a content publication platform (on the model of Medium), but with no data storage on a dedicated server. See also BAT (*Basic attention token*): https://basicattentiontoken.org/.

Moreover, it is not necessary to use this kind of technology to log and monetize data[7].

In our solution, the *smart contract* is used as a bridge between personal data, whether situated in a database or directly held by their owner.

The blockchain here is only the register of the transactions and is not used for storing data. Indeed, a transaction on a blockchain provides proof of publication[8], which avoids the duplication of a data item without the consent of its holder. In this context, transfer of responsibility is total. The data belongs to you[9] and the law guarantees ownership of them. If Bitcoin allows you to be your own bank, our solution is relatively comparable to the extent that you are now responsible for retaining and protecting your data. You decide or not to share them. Nevertheless, unlike Bitcoin, you have an ownership right guaranteed by the State.

Although it is now clear that blockchains provide proof of transfer and assignment of an individual's personal data to a third party, our "techno-legal" solution has at the present time certain limits that should be mentioned here.

7 Without necessarily monetizing data, researchers at MIT have developed the protocol HTTPa (standing for Accountable Hypertext Transfer Protocol), which logs the use of data from one server to another. Certain restrictions can be included upstream. See O. Seneviratne and L. Kagal, HTTPa: Accountable HTTP, November 2010, https://www.iab.org/wp-content/IAB-uploads/2011/03/oshani_seneviratne.pdf. Sweeney et al. propose for their part to associate with each transferred file, a system of *'datatags'*. This concept introduces in sharing a file, a *'tag'* which controls access depending on the degree of sensitivity of the data transferred. See: https://techscience.org/a/2015101601/. A more radical solution would be to move towards decentralized platforms, where all user data are stored on servers of their choice. See: https://www.w3.org/2008/09/msnws/papers/decentralization.pdf.

8 This term needs of course to be distinguished from the 'proof of work', which, in the Bitcoin and Ethereum protocols, is used to reach a consensus on the legitimacy of all the transactions.

9 We can already see this kind of solution emerging with distributed applications which work with Ethereum blockchains, such as for example IPFS (https://github.com/ipfs/ipfs).

The first limit is linked to the territoriality of the law, which may conflict with the fact that data are geographically agnostic. The second stems from the problem of identity on a blockchain that could cause certain legal complications. We will evaluate how decentralized market places can make monetizing data problematic (cf. Annex 2).

4. The socio-economic questions of a technological solution

• **Non-territoriality of data**

Beside the technical limits[10] of scaling-up the solution, the complexity of human interactions introduces specific problems relating to social sciences, even if they are not insurmountable over the long term.

In addition, the nature of the data, which can be easily copied, does sometimes render the maintenance of exclusive and non-rival property rights complex. We could go even further in this theoretical discussion: the location of a server and a database necessarily has an impact on the efficiency of an ownership right. Let us take a real example to illustrate our point. The publication of academic journals is not made directly by universities but by recognized publishers such as *Springer* or *Elsevier*. Researchers submit their articles to these publishers, then responsible for what is called “peer review”. The articles are sent anonymously to other researchers for comments. This process is organized by the publishers who, once this work is completed, publish the article. It is then sold by the publisher either per copy, or in the form of a subscription. Access to academic research is therefore limited to a few journals or platforms and is very expensive for the universities.

Challenging this monopoly of publishers, Alexandra Elbakyan, a Kazakh neuroscientist founded Sci-hub in 2011. The site today houses more than 64 million articles[11]. It has become the largest virtual library of academic articles. The site bypasses the journals’ payment systems using existing identifiers and directly downloads the unlocked article to a dedicated server. Most often, the content is retrieved illegally. However, the site

10 Scaled up, high transaction costs for nano-payments using protocols still under construction or that need refining. Scaling up public blockchains (https://blockgeeks.com/guides/blockchain-scalability/) is, currently, the greatest obstacle to implementing our solution, to the extent that we know in advance that a vast number of smart contracts will be triggered for the remuneration of personal data in the form of micro-payments. On Ethereum, work is in progress: https://plasma.io/plasma.pdf.

11 Source: https://en.wikipedia.org/wiki/Sci-Hub.

is so simple to handle that most researchers use it even if they have credentials provided by their own university[12].

The founder of the site is being sued by many publishers for violation of intellectual property rights[13]. Although several U.S. courts have already ruled that the site is illegal and have convicted its founder in absentia to substantial fines, Sci-hub continues to be used and to copy content originally protected by intellectual property. Elbakyan uses several domain names and the site is also accessible via *Tor*, a computer network which can be used to make the origin of connections disappear. Finally, the servers are not located on American territory and the illegal database cannot therefore be seized or destroyed.

How does that concern our data? **This example shows two limits to our solution. First, that data can be easily copied and stored without the consent of the counterparty. Secondly, that the law has territorial limits. If a malicious person decides to plunder personal data and is located geographically in a place where the rule of law is different, it will be very difficult to enforce it.**

The questions is: would the States that do not wish to apply this solution be able to enter into this kind of balance of forces? Currently, there is no technical solution to restrict data duplication.

- **Digital identity and physical identity**

A second limitation comes from the blockchains themselves. **The transaction system is fully pseudonymous**, if not anonymous. Let us take again the example of a current transaction. If I sell a property, the solicitor has the role of a trusted third party. The latter disappears if the transaction is now done through a self-performing contract. This poses no problem from the point of view of the exchange, provided the two parties have agreed on its terms. However, the solicitor also has the role of guarantor that the co-contractors are in full possession of their

12 Source: http://www.sciencemag.org/news/2016/04/whos-downloading-pirated-papers-everyone.
13 Source: http://www.sciencemag.org/news/2017/11/court-demands-search-engines-and-internet-service-providers-block-sci-hub?utm_source=sciencemagazine&utm_medium=facebook-text&utm_campaign=scihubblock-16248.

faculties at the time of the transaction. The same transaction via a self-performing contract on a blockchain therefore has two limitations. First, it is difficult to check that the private key has not been misappropriated and that the transaction is legitimate. Second, even if there has been no theft of this private key, nothing guarantees that I am not under guardianship for example, and that therefore I am in full possession of my faculties at the time of the exchange.

These difficulties related to identity are real and have a potential impact in legal terms. This directly relates to a problem of **repudiation of the contract** when a dissatisfied co-contractor will have a greater interest in not honoring their commitments. Since the identity is not formally verifiable, a person can always argue that someone has robbed them or that they have lost their private key. In addition, we have already shown in these pages the propensity of individuals to neglect their security on the Internet. The intermediary will probably be encouraged to facilitate the task of its customer by offering them to keep the private key and protect their access with a password. These two aspects have the effect of increasing instability and uncertainty related to the "contract".

Certain technological safeguards exist but currently remain imperfect and still need further research and development. **It may be possible to circumvent this problem in part thanks to the multi-signature. To be valid, a transaction must be signed by several parties to reduce the risks mentioned above.**

More generally, individuals should be able to self-manage their digital identity[14]. This is not the case today, to the extent that it has become a product used by platforms to maximize their advertising income. That is why defenders of self-sovereignty intend to give the control of his digital identity to the individual[15]. Applications existing or being developed already allow users to better protect ther data related to their identity[16]. These applications also provide part of the response to the limitations

14 Source: https://arxiv.org/pdf/1712.01767.pdf.
15 Vitalik Buterin briefly decries these points in a recent interview: http://unchainedpodcast.co/.
16 We are thinking here of uPort: https://whitepaper.uport.me/uPort_whitepaper_DRAFT20170221.pdf, or CIVIC: https://www.civic.com/products/how-it-works.

related to the identity theft mentioned above. These techniques are used to formally identify each user.

- **Granting of rights vs control**

One of the aspects developed in our analysis relies particularly on the concept of control of data. This is a central theme because the consumer-citizen needs to be put back at the heart of our solution, giving them back a certain amount of control over the use or non-use of their data by a third party. The very concept of control, from a twofold technical and legal aspect can sometimes be ambiguous, depending on whether you place yourself on the side of the technical or legal point of view[17]. Although we have seen the legal outlines of the concept of control of personal data, we need to say a few words about its technical aspects.

As a solution for the decentralized production of proof, blockchains could be a reference registry, registering not the transactions of the data themselves but the granting of rights to use these personal data. **One could imagine a market of rights relating to the use of the data. A kind of on-air broadcasting market applied to the use of personal data.** The duration and the scope of the use of these data could take the form of metadata embedded in the transactions. The transactions recorded in the blockchains could constitute proof of granting of use of the data and their terms of use. This could constitute legal proof of the use of personal data. The citizen could invoke this transaction in the event of a dispute. It is this reference, this proof that a right has been granted, which may increase in value and therefore take the form of a traded asset, rather than the data themselves.

17 The difficulty of control over personal data, from a technical point of view anyway, is particularly true in the era of Big Data, see for example: https://scripted.org/article/control-over-personal-data-true-remedy-or-fairy-tale/. It is specifically for this reason that a technological solution must be associated with a legal guarantee, see WERBACH K., "Trust purpose verify: Why the Blockchain Needs the Law?" Berkeley Technology Law Journal, 2018.

The challenge is to build a market of data use rights before creating a data market.

06

CONCLUSION

We believe in a decentralized Internet, where individual identity still has meaning. The individual is not simply reduced to a flow of data. This spirit has influenced our proposals.

The advent of the digital age in the early 2000s brought the hope of a certain degree of decentralization of power as well as free flow of information. In reality, we are experiencing a concentration of power in the hands of those who have unlimited access to the information that we produce every day through our data.

We believe in a decentralized Internet, where individual identity still has meaning. The individual is not simply reduced to a flow of data. This spirit have influenced our proposals, which are structured around two focus areas: one is technical and the other legal so that each person can regain control of his digital identity.

Pushing forward the public debate on these issues is the ambition of this report, which unfortunately cannot cover all aspects of such huge and complex questions. If the analysis set out here is mainly a legal one, a second publication will focus on the economic aspect and the value given to personal data.

For now, and as a conclusion, the technicality of some of our proposals may be reflected through a few possible future scenarios.

9 September 2019. Gérard only drives on Sundays to visit his daughter and grandchildren. Dozens of kilometers separate them and Gérard is very familiar with the road. He nevertheless uses a real-time traffic application to avoid congestion on the Paris ring-road. For her part, Camille drives her truck all over France for a transport company and also uses the same application.

In both cases, access to the application “costs” the same thing (i.e. nothing) to a truck driver who drives every day or to a retiree who only drives his car on Sundays. In exchange, they both have to share their geolocation to inform the application about the status of the traffic. Camille therefore contributes much more than Gérard to the value of the application.

If data ownership came about, Camille could demand to be paid, since the value that she produces for the application is much greater

than what she receives from it. In contrast, Gérard could choose to pay for access to the application without transferring his data, thus compensating it for his 'stowaway' behavior. In this way, everyone, depending on their life style and uses of the Internet, everyone could build the economic model which best suits them[1].

21 June 2021. Karim struggles to wake up. Summer light floods into the room as the blinds in his apartment rise automatically. A radio program is activated on his smartphone. Karim is hungry. But he ignored the warning from his refrigerator about the shortage of food needed for breakfast and drinks coffee already prepared by his machine connected to the alarm on his phone. In his home, everything is programmed remotely and interconnected to make sure he wakes up in the best possible way. Everything is calculated based on his sleeping pattern and the time he needs before going to work.

The constant interconnection between different devices generates a significant amount of data about his daily comings and goings, his health, and consumption habits[2]. These devices are also a source of IT risks. Karim is aware of it and wants to protect his intimacy, while limiting any potential piracy of his home automation.

In this case, two options are available. The first is to select each of the devices that include in their design a mechanism for data protection, in specifically using encryption code[3]. However, the transfer of data between the different devices is not always guaranteed. Karim opts for a distributed solution. Each of the devices is connected to the same API. The data are transferred via a dedicated blockchain[4], which protects Karim's home automation system from potential attacks while at the same time guaranteeing the confidentiality of his private life.

1 The Steamr application takes a step in this direction: https://www.streamr.com/#howItWorks.
2 About the risks of the Internet of Things and the end of confidentiality, see for example: https://dl.acm.org/citation.cfm?id=3105843&dl=ACM&coll=DL&CFID=850514586&CFTOKEN=59441772.
3 This is what the English-speaking world calls *privacy by design*.
4 We are thinking here of IOTA: https://iota.org/IOTA_Whitepaper.pdf. Although this blockchain is criticized (https://hackernoon.com/why-i-find-iota-deeply-alarming-934f1908194b), the idea of a blockchain dedicated to the Internet of Things remains of interest.

22 January 2022. Alice is 19 and wants to open her first ever an account on a social media platform. Instead of electronically signing a "user agreement" that she will not read, a home page opens, listing all the data the platform will offer to pay in exchange for their use by the company.

Two prices are displayed for each type of data (age, ethnicity, city, analysis of photos, political preferences, etc.)[5]. The first is the purchase price by the platform of Alice's data, the second is the price for the protection of her personal data. These prices have been calculated based on how she previously replied to a few questions about her browsing habits.

Now an adult, Alice has been proactive and has already transferred the data that she agrees to share on the social media to her data broker. The broker keeps them on protected servers, set up in Alice's country of origin. The only thing she had to do was to provide the name of her broker and her contract number[6] on the dedicated page of her new social media account. This contract is triggered automatically every time all the pre-established conditions are met. For example, if Alice has agreed to share her age and gender with the platform, she will be paid annually for their use via a smart contract. Validation by Alice via her contract number is what triggers the *smart contract*.

Each year, her broker will send a statement of what she owes the platform to protect her data. If the platform owes her money, it will pay her the corresponding amount in euros. Alternatively, Alice could choose to manage her data alone and to protect them on a dedicated IT medium. In this case, she will have to inform the social media of which data she agrees to make public and which ones she doesn't. A contract will then be directly created with the social media platform.

5 This reflects the latest work on behavioral economics, where the price that a person is willing to pay to protect their data is often different from the price that they are prepared to receive to transfer them. See ACQUISTI A., BRANDIMARTE, L., LOEXENSTEIN G., *Op. cit.* And https://www.cmu.edu/dietrich/sds/docs/loewenstein/WhatPrivacyWorth.pdf. This price strategy could possibly be taken into account by the platforms to reflect this cognitive bias.
6 When the conditions of use change, the *smart contract* then becomes obsolete. This has a twofold effect. First, it encourages the supplier of a service not to change them too regularly, otherwise it risks seeing access to its service reduced during a transition phase. Second, we see here a certain limitation to the *smart contract*, which is fairly rigid and does not change in line with the updates of a non-paying site.

Alice now pays for her service with her personal information. In both cases, validation of the transaction via a blockchain is used to authenticate a transfer of rights to these data, and to be paid accordingly.

07

ANNEXES

Technical details

Annex 1. Analysis and evaluation of the data

Whether it is from the point of view of the government or large Internet companies, personal data play an increasingly important role in the analysis of behaviors. Most often, it is not your individual data items in themselves that are of interest and are monetized, but their aggregation[1].

Consequently, a change in the business model implies a change in the value chain. We leave this issue pending in this report, focusing primarily on the legal aspect of the data and tries to answer the following question: how can the consumer being given power over his data ? Although the economic approach is of real interest to initiate a debate, this overall technical issue is too broad for this publication and will be addressed subsequently. The question is already thorny, to the extent that it involves issues of public policy that lawmakers only start to deal with now.

The objective is to paint a broad outline of the economic issues and give the reader a brief overview of the literature[2].

At the macroeconomic level, allocating an intellectual property right can be used to eliminate Arrow's information paradox set out in his 1962 paper. Let's suppose that I want to sell the data linked to my identity to a third party. During this transaction, the buyer will know

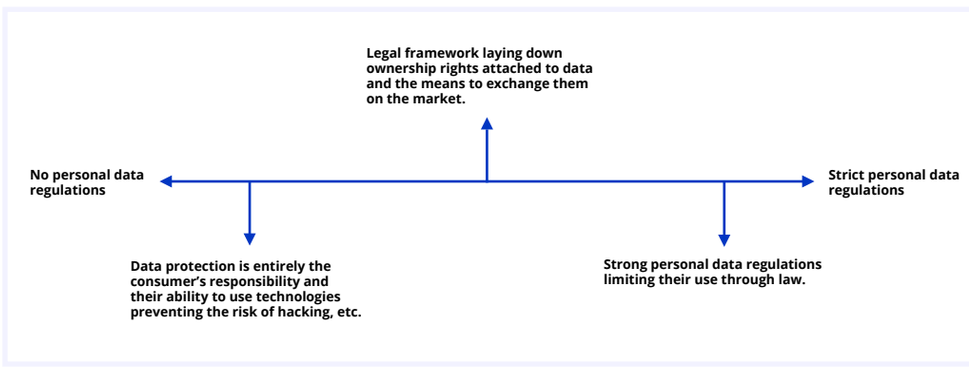
1 It should be noted that economies of scale cannot always be made. The increase in the number of observations does not necessarily improve the return on investment. This is the case for example in advertising. Lewis, R. and Rao J., *The unfavorable economics of measuring the returns to advertising*, Quarterly Journal of Economics, 130(4), 2015.

2 Literature reviews are available, see in particular: ACQUISTI A., TAYLOR C., and WAGMAN L., *Op. cit.* or: https://ec.europa.eu/jrc/sites/jrcsh/files/jrc104756.pdf.

how much he is prepared to pay for this information. This is called in microeconomics terms "his propensity to pay". However, to make this assessment, the buyer can ask the seller to reveal information so he can determine the right price. As opposed to a good or service, the data disclosed no longer have any value. In this context, intellectual property appears to be a natural candidate for the protection of personal data.

Although it seems that the Arrow paradox can be overcome, there is another problem related to the marketing of data: they are non-rivalrous and non-excludable[3]. First, this means that the information can be copied without affecting its use by another consumer. Second, it is difficult to exclude access to this information by the user, even via the intermediary of a price system. Indeed, a buyer may choose to resell the data that it has just acquired, if the law allows it.

An approach in terms of property rights can be used to regulate the exchange of data rather strictly[4], as illustrated in the following graph:



On the other hand, this kind of model would lead to reducing the consumer's surplus[5]. There is therefore a cost/benefit to the excludability

3 Source: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2858171.

4 In relation to this, Europe and the United States have different approaches. As we have seen, the GDPR opens the door to assigning ownership to data. Whereas the United States defends a position which facilitates the appropriation of personal data by *databrokers*. This market is little known to consumers, to such an extent that the *Federal Trade Commission* is calling for greater transparency and improved oversight. Source: https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

5 ACQUISTI A., TAYLOR C., and WAGMAN L., *Op. cit.*

that must be assessed. Before being able to calculate this cost, it is necessary to establish a market price for the data.

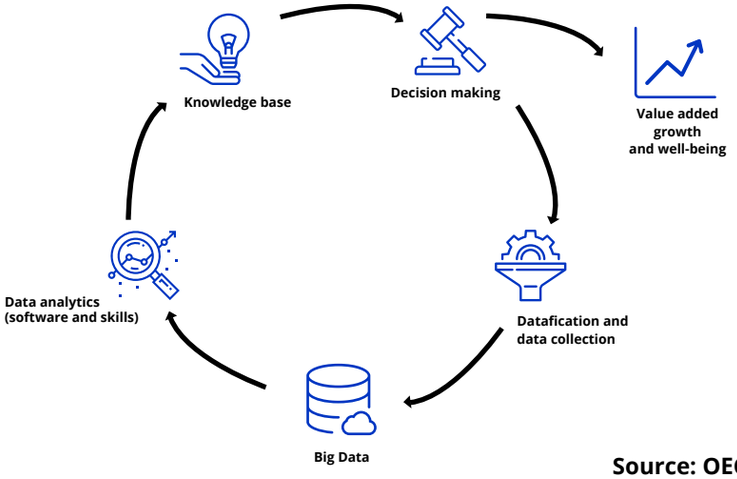
Moreover, collection and reprocessing of the data represent a cost for the organizations that they must assess[6]. This assessment is necessary because it is central to our knowledge economies, where value is derived from information.

First, the new information and telecommunications technologies have economic models that work based on these data[7]. They become the raw material of the new digital economy. The dematerialization of our relations gave a way to a new kind of benevolent "surveillance". If information has an economic value, we can be sure that it implies a new balance of power.

Second, the analysis of the data is itself part of a value chain, which can have a positive impact on growth. The following graph below illustrates this cycle of the value chain:

6 We refer the reader directly to the work of the OECD on this subject, which provide a better understanding of the difficulties that are met in this complex exercise. Source: https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP/IE/REG(2011)2/FINAL&docLanguage=EN and http://predipubcn.sistemaip.net:8096/intranet-tmpl/prog/img/local_repository/koha_upload/DSTI-CDEP(2016)4-ENG.pdf.

7 Varian H., Farrell J., and Shapiro C., *The economics of Information Technology: An Introduction*, Cambridge University Press, 2004.



Source: OECD[8]

Third, the platforms are a rather unique market mechanism. Since Rochet and Tirole's work[9], we know that these two sided markets must satisfy customers with heterogeneous profiles: on the one hand, the users, on the other, the advertisers. In the first models, the price mechanisms were based on a network effect, and not on the collection and data use. These models today are evolving to better integrate the effects on well-being as the platforms would reduce transaction and information costs[10].

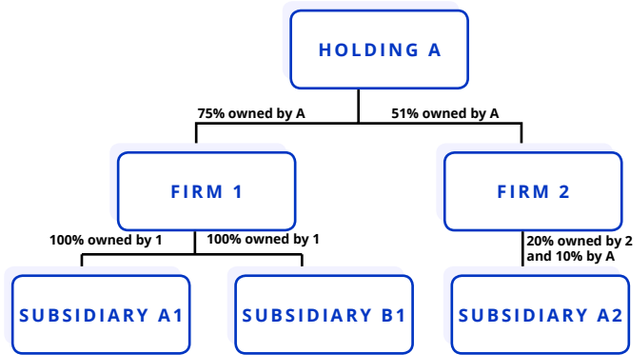
Finally, there is an obvious behavioral component that influences individual choices and highlights the difficulty of establishing a price for data, which depends mainly on the context[11].

8 Source: *Ibid.*
9 Rochet J. C., and Tirole J., *Two-sided markets: a progress report*, Rand Journal of Economics, 37(3), 2006.
10 Source: https://ec.europa.eu/jrc/sites/jrcsh/files/jrc104756.pdf.
11 ACQUISTI A., BRANDIMARTE, L., LOEXENSTEIN G., *Op. cit.*

Annex 2. The case of decentralized "Data Market places"

A data market is **an information market**. Would allocating **property rights** help correct some **shortcomings in the market**? One of the particularities of the information economy was set out in 1962 by the economist Arrow[12]. Sharing of information introduces this paradox.

Let us consider this very simplified ownership structure; in general, large groups have tens or even hundreds of subsidiaries that they control more or less directly. A holding company A has a majority shareholding in several companies. Let us suppose that all these companies are recognized as "establishments" by the new European regulations and that they are all located in the EU. This hypothesis removes certain difficulties and provisions specific to the GDPR and related to the transfer of data from the EU to third countries. In this case, such a cascading structure would allow a holding company to share its customers databases with several subsidiaries, without the customers having necessarily given their consent, in this case with firms 1 and 2 and their subsidiaries A1, and A2. This is due to the ownership structure in our example. In this context, the data partly escape from the control of their holder. This results in an additional difficulty of control for the regulator. In this case the regulations should provide for the establishment of a sufficiently transparent system to enable the user to easily know where their data are.



12 Source: http://www.nber.org/chapters/c2144.pdf.

Although a legal contract can specify in detail what entity holds what data, this also generates **monitoring** costs that are difficult to dissociate from the said contract. Indeed, it would be much too expensive to store each data item related to each transaction on a blockchain. This new architecture is not designed for this. In this context, monitoring compliance with the law will have to be done on a case by case basis, where data are indeed exchanged, within each company. The information needed by the court to apply the law should therefore be easily verifiable by this third party. However, the court is itself subject to two constraints well known to economists: moral hazard and adverse selection. The first assumes that the judge has not made the necessary effort to understand the ins and outs of the case. The second stipulates that the judge does not have the necessary knowledge and/or the information to deal with this case in all fairness or that (s)he has his or her own preferences in terms of case-law. These information asymmetries are reinforced by the fact that the technologies used here are new and not yet mastered by law professionals. On this last point, some jurists talk about the "ossification of the law"[13]. These inherent uncertainties in the judicial decision result in an additional risk for the stakeholders, namely that neither the letter nor the spirit of the contract are implemented[14].

Let's imagine an exchange of digital personal data platform, on which we could put up for sale the data of our choice and find an interested buyer. The data transactions would be recorded on the register of a public blockchain in return for remuneration in crypto-currency. The only problem is that once the data are exchanged, nothing can prevent them from being duplicated. A single exchange is enough, in theory, to destroy your own market. And although it contains proof of the transaction, a blockchain seems here powerless in addressing the risk of "infringement".

13 MC GARITY, Thomas O. 1992. Some thoughts on deossifying the rulemaking process. Duke Law Journal 41: 1385, 1385–1462 ; Pierce, Richard J. Jr., 1995. Seven ways to deossify agency rulemaking. Administrative Law Review 47: 59, 60.

14 Source: https://www.jstor.org/stable/2999457?seq=1#page_scan_tab_contents.

08

Main references.

Books.

• BENYAYER L-D., CHIGNARD S., *Datanomics, les nouveaux business models des données*, FYP Editions, 2015.

• MANION Josh, *The power of data ownership : getting it right in 2017*, December 20, 2016.

• LOSHIN David, *Business-Oriented Data Governance for effective Master data Management*, 2015.

• LANIER J., *Who owns the future ?*, Simon and Schuster, 2013.

• LUCAS André et LUCAS Henri-Jacques, *Traité de la propriété littéraire et artistique*, Litec, 2012.

• SIMLER C., *Droit d'auteur et droit commun des biens*, Litec, 2008.

• WRIGHT D., DE HERT P., « Enforcing Privacy Regulatory, Legal and Technological Approaches. Law », *Governance and Technology Series*, Volume 25, 2016.

• SFADJ Rubin et GRANGER Elodie, *Réussir votre mise en conformité GDPF : Guide pratique*, Broché, 2017

Reports.

• ARROW K., *Economic Welfare and the Allocation of Resources for Invention*, The Rand Corporation, 1962.

• Organization for Economic Co-operation and Development (OECD), *Thirty years after the OECD Privacy Guidelines*, 2011.

• OECD, *Exploring the economics of personal data : a survey of methodologies for measuring monetary value*, 2013.

• SEN A., STIGLITZ J., FITOUSSI J-P, Rapport de la Commission sur la mesure des performances économiques et du progrès social, La Documentation française, sept. 2009.

Articles.

• DASKAL Jennifer, "Beware of the Emergency Exception Loophole in the Email Privacy Act", Just Security, June 7, 2016.

• NAKASHIMA Ellen, "FBI Wants Access to Internet Browser History without a Warrant in Terrorism and Spy Cases", Wash. Post, June 6, 2016.

• GREENE Robyn, "FBI's Push to "Fix a Typo" Would Really Expand Its Surveillance Authority", Just Security, February 17, 2016.

• TRUJILLO Mario, "California Puts Email Privacy Law on the Books", The Hill, October 9, 2015.

• TRUJILLO Mario and McCABE David, "Overnight Tech: Senate Gets Late Start on Email Privacy", The Hill, Sept. 15, 2015.

• Sen. LEAHY Patrick & Sen. LEE Mike, "Update Privacy Laws for the Digital Age", Real Clear Policy, Jan. 28, 2015.

• FARIVAR Cyrus, "Unprecedented E-mail Privacy Bill Sent to Texas Governor's Desk", ArsTechnica, May 28, 2013.

• Editorial, "Upgrade Protections of Digital Records", Seattle Times, May 16, 2013.

• McGREEVY Patrick, "California Senate Backs Requiring Warrants When Police Want E-mails", L.A. Times, May 13, 2013.

• ACQUISTI A., TAYLOR C., WAGMAN L., « The Economics of Privacy », *Journal of Economic Literature*, 54(2), 442–492, 2016.

• CANTERO Isabelle, « En attendant le règlement européen sur les données personnelles… 'l'essentiel fait vertu' », *L'Usine Digitale*, 19 février 2016.

• LECHENET Alexandre, « Données de santé : une base saine mais peut mieux faire », *Libération*, 5 mai 2016.

• HERAULT S., BELVAUX B., « Privacy paradox et adoption de technologies intrusives. Le cas de la géolocalisation mobile », *Décisions Marketing*, 74, 2014.

• MATTATIA F., YAÏCHE M., « Etre propriétaire de ses données personnelles : peut-on recourir au régime traditionnel de propriété ? », *Revue Lamy de droit immatériel*, 114, 2015, p.62.

• NAKAMOTO S., Bitcoin : « A Peer-to-Peer Electronic Cash System », 2008 : https://bitcoin.org/bitcoin.pdf

• ROCHET J. C., TIROLE J., « Two-sided markets: a progress report », *Rand Journal of Economics*, 37(3), 2006.

• ROUX D., « La résistance du consommateur : proposition d'un cadre d'analyse », Recherche et Applications en Marketing, 22, 4, 2007, pp.59-80.

• Science, *The End of Privacy,* vol . 347 (6221), 2015.

• WRIGHT P., « Marketplace Metacognition and Social Intelligence », Journal of Consumer Research, 28, 4, 2002, pp. 677-83.

Websites.

• Une carte interactive de la CNIL pour savoir où vos données vont être exportées : https://www.cnil.fr/fr/la-protection-des-donnees-dans-le-monde

• Ethereum White Paper : https://github.com/ethereum/wiki/wiki/White-Paper

Conferences.

• #Dataday, Conférence débat sur la stratégie d'open Data pour le développement de l'économie de la donnée , 12 janvier 2016.

09

ACKNOWLEDGMENTS

Authors

Nicolas BINCTIN

Law Professor, Nicolas Binctin teaches the various components of business law and intellectual property law at the Universities of Poitiers and Paris XII, and the School of Law and Management of the University of Paris II.

Isabelle LANDREAU

Attorney at the Bar of Paris and Doctor of Law, Isabelle Landreau specializes in intellectual property law and new technologies law. She assists her clients in protecting and enhancing their immaterial creations and practices as a lawyer-mediator in intellectual property.

Gérard PELIKS

Engineer in cyber security, Gérard Peliks has worked in the information security field for over 20 years. President of CyberEdu, he is also Deputy Director of the MBA Management of the security of Digital Data at the "Institut Léonard de Vinci".

Virginie PEZ-PERARD

Teacher-researcher, Lecturer at the University of Paris II Panthéon-Assas, Virginia PEZ is a specialist in consumer behavior, the psychology of consumption and issues of privacy and intrusiveness in commercial practices.

10

GENERATION LIBRE

The Think Tank's Purpose

Tocqueville already deplored in *The Old Regime and the Revolution*, "the frightening spectacle" of French philosophers, cut off from their peers, unaware of the life of society and blind to the rest of the world. "The same attraction for general theories, comprehensive systems of legislation and exact symmetry in the laws; the same contempt for existing facts; the same confidence in theory."

Conversely, politicians are too often detached from any philosophical reflection, relying too much on the civil service to come up with reform projects.

"It is for the purpose of a better combination between theory and practice, philosophical principles and political action, that think tanks must work"

On the basis of a clear doctrine, they bring together the skills of experts to sometime shape unusual ideas into specific and quantified public policies. With regard to universal basic income for example, GenerationLibre took hold of a powerful but very abstract concept and developed an economically viable proposition in the form of a negative tax.

It is positive that think tanks are playing an increasingly large role in the French public arena. Beyond the convictions of each person, this is the guarantee of a rich and informed debate on the major issues of our age.

ACTIONS

Our daily challenge

Our objectives

1. **Live and let live,** to allow everyone to define their own values in an open society.

2. **Liberalize the economy,** because the free trade of goods and services like ideas is the best way to challenge the established order.

3. **Apprehend progress,** so that technological innovation continues to benefit individuals.

Our latest publications

- "Redesigning employment contracts: from subordination to cooperation", January 2017;
- "LIBER, a realistic proposal, tome II", January 2017;
- "Gender and the State: from the status of non-disposal to free determination", June 2017;
- "Refounding Europe, for a minimal European State", Chapter I, April 2017;
- "Schumpeter and robots, the case of France", November 2017.

SUPPORT US

Support new ideas

founded in 2013 by the French philosopher Gaspard Koenig, GenerationLibre is a liberal think tank. It is financed exclusively by the generosity of its members, the only guarantee of its freedom and its independence. It refuses any public subsidy and is not involved in any consulting activity. GenerationLibre pursues 3 objectives: to live and let live, to liberalize the economy, and to apprehend progress.

Write to us, come and see us

GenerationLibre
24, rue Saint-Lazare
75009 Paris
contact@generationlibre.eu

www.generationlibre.eu