INSTITUT
MONTAIGNE

Digital Privacy:
How Can We Win the Battle?

STUDY NOVEMBER 2019

There is no desire more natural
than the desire for knowledge

INSTITUT
MONTAIGNE

Digital Privacy: How Can We Win the Battle?

NOVEMBER 2019

About the author

François Godement, Senior Advisor for Asia, Institut Montaigne

François Godement is Senior Advisor for Asia to Institut Montaigne, Paris. He is also a non-resident senior associate of the Carnegie Endowment for International Peace in Washington, D.C., and an external consultant for the Policy Planning Staff of the French Ministry of Foreign Affairs. Until December 2018, he was the Director of ECFR's Asia & China program and a Senior Policy Fellow at ECFR. A long-time professor at France's National Institute of Oriental Languages and Civilisations and Sciences Po, he created Centre Asie IFRI at the Paris-based Institut Français des Relations Internationales (1985-2005), and in 2005 Asia Centre as an independent. He is a graduate of the Ecole Normale Supérieure de la Rue d'Ulm (Paris), where he majored in history, and a postgraduate student at Harvard University. In 1995 he co-founded the European committee of the Council for Security Cooperation in the Asia-Pacific (CSCAP), which he co-chaired until 2008. He has also been a member of the advisory board for the Europe China Academic Network (ECAN).

SUMMARY

INTRODUCTION

"Gentlemen don't read other people's mail." Actually, they sometimes do, legally or surreptitiously. The novelty is that they no longer need to steam open an envelope. In fact, we ourselves are emitting personal data around the clock rather than write old-fashioned letters. Barring end-to-end encryption, or more rarely content encryption, that data floats in cyberspace.

But how can gentlemen find their way in the exaoctets of data flowing through cyberspace from the sanctuary of your home, or among the nearly 3 billion smartphones in circulation, and tomorrow the zillions connected objects? Isn't there safety in numbers? Most digital users do not trust their data to be completely safe, but they rely on a degraded version of trust – nobody will look for a pin in the haystack.

With this reservation in mind, the digital age has become the last free frontier in our age. To use a metaphor, the digital age is to its predecessors what maritime exploration was to bound land states, the "21st century equivalent of the 'dark continents' that drew 19th century European speculators to their shores."[1] Or perhaps, what the opening of the American West represented to its pioneers from original Eastern American states: lands of opportunity where the laws hardly applied. With its drawbacks: corsairs – "privateers" were an official practice for maritime nations. One was free at sea, including to fall prey to corsairs or marauders. Similarly, the "law of the West" was a euphemism. Social rumors or fake news, "scraping" for commercial use of everyone's "small data", cybersecurity issues, are the contemporary equivalent.

[1] Shoshana Zuboff, *The Age of Surveillance Capitalism : The Fight for a Human Future at the New Frontier of Power* (New York: Publicaffairs, 2019), p. 103.

Living on the Digital Frontier

The digital revolution – a combination of big data,[2] predictive algorithms and quantum computation – promises to change our lives as no other breakthrough ever did. This is because connecting numbers with one another, even as it concerns "things" as in "the internet of things", in fact goes to the heart of human behavior, and one day to our inner thought process. For the time being, the revolution is satisfied with statistical predictions based on digging into deep layers of human behavior, but one day these predictions, and the behavior variables that they are based on, will become so granular and accurate as to be contemporary or even synonymous with human thought processes. The combined performance of algorithms and big data does not stop there. Machines have beaten human beings at chess, at go – a much more multi-dimensional experience, and now at poker.[3] Artificial Intelligence (AI) interpretation of lung imagery for tumors and choice of treatments consistently beat the best human medical teams.

Along with those achievements come visions of utopia. Who needs physical libraries if internet clouds provide much more capacities – and major libraries too become clouds: the Library of Congress, for example, has received in deposit a complete archive of ALL tweets posted on Twitter for the first decades. The average smartphone gives access to a bandwidth of information that was never accessible in any form to any individual in the previous era. There are those who argue that the digital revolution, complemented by edge (local)

[2] For a clear introduction and historical account of this: Gilles Babinet, *Big Data, Penser l'homme et Le Monde Autrement* (Paris: Le Passeur Éditeur, 2016), p. 25-51.

[3] Noam Brown and Tuomas Sandholm, 'Superhuman AI for multiplayer poker', *Science,* August 30, 2019.

computing and servers, and software incorporating AI will render obsolete the Fordist and Taylorist industrial revolution. Imaging, remote diagnosis and predictive diagnosis will revolutionize preventive medicine, and allow for the treatment of billions of people who had no suitable access to doctors and medical resources. Automated decisions will become commonplace, shrinking the drudge of daily chores, just as physical work has shrunk in the course of the first industrial revolutions. In the end, we will be pure minds, focused on innovation, leisure and instant communications with all the other monads in the world, overcoming physical, language and cultural barriers.

Or will we?

Brave New Data World

The other vision is one of dystopia. Some of it is economic, with the prospect of mass unemployment: white-collar jobs, including many that were previously thought as skilled, will be automated. But the prevailing dystopian view focuses instead on the disappearance of privacy for marketing or control purposes. The case for this is almost the same as that for the utopian vision, so much so that one may consider it as its flip side.

The conflict between individual and sovereign rights is the oldest conflict in political philosophy. But in the digital age, one can replace individualism with the issue of privacy. There can be no individual right if there is no privacy. If polling becomes so granular as to predict the vote of a given individual with near certainty, there is no longer any confidentiality of voting. The Obama 2008 presidential campaign

compiled data on more than 250 million Americans. In one participant's words, "we knew… who people were going to vote for before they decided." The president's 2012 re-election campaign knew "every single wavering voter that it needed to persuade to vote for Obama by name, address, race, sex and income", and created "persuasion scores" for undecided voters.[4] The Obama campaign broke no law that we know of, and it has in fact become a model for campaigning elsewhere in open democratic contexts. Some of the same operators would work later for the Trump campaign and for Cambridge Analytica, probably the company involved in the most glaring privacy scandal of recent times.

"With near certainty": in that tiny islet where chance can still prevail over necessity, can we really find the remaining element of human choice and self-determination? That is only conceivable in a society which is protected by positive law and institutions. Elsewhere, to categorize or to judge you, near certainty is good enough.

Predictive Chaos

The heart of the legal debate about individual rights – *habeas corpus* – has often been about the legality of practices, not about their reality that made little doubt: habeas corpus itself means that the "body" or subject must be brought to justice, and not disposed otherwise. In terms of digital data, it is about the legal duty to prevent "intrusion into seclusion", and about the public and legally admissible use of the data collected. Chinese fintechs run on big platforms with access to many more types of data than is legally permissible in the West:

[4] Cited by Shoshana Zuboff, *The Age of Surveillance Capitalism : The Fight for a Human Future at the New Frontier of Power* (New York: Publicaffairs, 2019), p. 122-123.

they can deliver a credit or insurance rating within seconds for the smallest of entrepreneurs, based on a swath of data that includes many personal habits and possible incidents. It is a near certainty that one individual or entity owning AI resources can game markets in the near future (one can postulate the same outcome for war scenarios). However, *if several* individuals or entities compete with algorithms, the result may well be chaos and complete uncertainty. The market's resiliency came from the fact, as suggested by Friedrich Hayek, that no individual human mind could outguess "the coordinated utilization of resources based on equally divided knowledge."[5] That is over. But even Hayek had not envisioned that algorithms could play each other, creating overall instability. The principle of uncertainty is not equivalent to market play.

The above should be nuanced with two interrelated issues: that of accuracy, closely linked to the quality of the algorithms used. Few areas of AI regarding human behavior are likely to offer the reliability that DNA typing (but not the handling and conservation of samples) has acquired. And big data banks are only as good as the software to interpret them. But that in turn should be relativized by another fact: relative certainty can be good enough in some systems. Facial recognition has become so commonplace there is an over-the-counter Russian app that makes it available to any web user.[6] But it is still riddled with "false positives" and "false negatives" (wrong identification or missing an actual match). Most experts recognize that the path from 90% to 100% accuracy is much harder than that from 50% to 90%, not including possible deception and concealment. Yet some

[5] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: Publicaffairs, 2019), p. 497.

[6] Kevin Webb, "Viral app that makes you look old with shocking precision may be quietly keeping all your data", *Business Insider France,* July 17, 2019, https://www.businessinsider.fr/us/faceapp-privacy-data-terms-service-russia-2019-7.

governance systems will be satisfied with 90% accuracy, and even more so if it is compounded with other predictive results.

The makers of the digital age are not the initial inventors of computing, or even the internet. They are the entrepreneurs who have turned these into a nearly universal commodity, and who have created an immense new field of social media and data use. They were often convinced that they were bringing unbounded freedom to individuals, especially from cumbersome regulations, and with this freedom, a blossoming of individuals. "The online world is not truly bound by terrestrial laws… It's the world's largest ungoverned space."[7] To all practical purposes, users in the digital age act with an illusion of privacy that is greater than at any other time: in fact, they consign most of their private data, one way or another, to the digital space.

Isn't it telling that meeting online has become the most popular way (39%) couples form, displacing the role that family, friends and public places once played?[8] It can be argued that this method of seeking partners offers more privacy than earlier methods involving intermediaries or public searches. Googling, roaming, exchanging over social media appears to enhance the individual against the constraints and inhibitions of the community. Cyberspace is both the largest public space ever and yet, it is thought to be a very private meeting place. If this wasn't the case, 30% of internet traffic wouldn't

[7] Eric Schmidt and Jared Cohen, *The New Digital Age Reshaping the Future of People, Nations and Business* (London Murray, 2014).

[8] 39% applies to heterosexual couples. For homosexual couples, the proportion of on-line meeting rises to 65% according to the same study.

Source: Michael J. Rosenfeld, Reuben J. Thomas, and Sonia Hausen, "Disintermediating Your Friends: How Online Dating in the United States Displaces Other Ways of Meeting," *Proceedings of the National Academy of Sciences* 116, no. 36 (August 20, 2019), p. 4, https://doi.org/10.1073/pnas.1908630116.

be about adult content, as has been the case until the new streaming media provided alternative distraction at home, taking 60 % of the overall internet traffic. This fact should come with the remark that in all likelihood, the immense majority of customers for this would not have dreamt of entrusting this type of content to the post office, even with gentlemen looking the other way.

There is no denying this new freedom, and the ample opportunities that have come with the digital age. But the other side of the coin has become increasingly clear: the "scraping" of the individual data we leave on the digital media, the very extent to which every one of our movements, actions and increasingly our thoughts is enacted through a digital medium and therefore open to scrutiny – perhaps forever and without reprieve – create a world of transparency, where surveillance is the practice if not the norm. Quite simply, we embrace new digital tools and social media platforms which make us less private. How to find a balance between our freedom and our privacy is an extremely hard choice, even at the individual level.

The Digital Age, Like the Nuclear Age, Cannot Be De-Invented

Big data is not only "the new oil." Algorithms make it the human equivalent of atom fission, otherwise known as the nuclear bomb. In Eric Schmidt's words, "almost nothing, short of a biological virus, can scale as quickly, efficiently and aggressively as these technology platforms, and this makes the people who build, control and use them powerful too."[9] Or as Jim Balsillie, ex-RIM CEO explains, "data

[9] Eric Schmidt and Jared Cohen, *The New Digital Age Reshaping the Future of People, Nations and Business* (London Murray, 2014), p. 9-10.

at the micro-personal level gives technology unprecedented power to influence. Data is not the new oil – it's the new plutonium. Amazingly powerful, dangerous when it spreads, difficult to clean up and with serious consequences when improperly used."[10] In the near future, the expansion of the digital age with the so-called Internet of Things (IoT) will place us in a web of myriad interconnected devices endowed with a form of machine intelligence, served with sensors registering their – and our – environment. To an extent, we will be the subjects driving these networks – or some operator will be that subject. But there is a very big likelihood that these networks will move "from a thing that we have to a thing that has us." Already, some of the most comprehensive owners and sellers of our personal data are companies whose names we don't even know. One third-party data broker, Acxiom (now renamed LiveRamp), claims to have amassed by 2018 up to 10,000 attributes on 2,5 billion individuals, a "comprehensive representation of 68 percent of the world's online population."[11]

Linking up very few metadata points from dispersed sources can now lead to identification of individuals. That holds true even when the collection of each of these data points has been anonymous. AI is a zillion minds put together, and it will move faster than any human thought and action in any case. A similar outcome can befall encryption, defeated by quantum computing. What matters with the

[10] Financial Post, "Jim Balsillie : 'Data Is Not the New Oil – It's the New Plutonium,'" *Financial Post,* May 28, 2019, https://business.financialpost.com/technology/jim-balsillie-data-is-not-the-new-oil-its-the-new-plutonium.

[11] The company has changed ownership after the 2018 Cambridge Analytica scandal. Source: Alex Pasternack, "Here Are the Data Brokers Quietly Buying and Selling Your Personal Information," *Fast Company,* March 2, 2019, https://www.fastcompany.com/90310803/here-are-the-data-brokers-quietly-buying-and-selling-your-personal-information.

above is not only that privacy, defined as confidentiality and a principle of uncertainty for human actions, is replaced with nearly complete knowledge of the individual mind, or at least categorization and predictability, it is also that there is a huge asymmetry created between the operators of the systems and their object – the individual. The prophets of this age have entirely anticipated this dispossession. B.F. Skinner, a leading behavioral psychologist, wrote in 1971: "*to man qua man* we readily say good riddance. Only by dispossessing him can we turn…from the inferred to the observed, from the miraculous to the natural, from the inaccessible to the manipulable."[12]

The dispossession extends to uncontrollable human features – deducing human preferences and feelings. "With enough big data, the numbers speak for themselves."[13] From Facebook likes, sexual orientation can be deduced with 88% accuracy. From keyboard typing patterns, sadness can also be detected with 88% accuracy: these achievements date back to 2011, and one can only guess the progress made since then.[14] And of course, asymmetries exist also between operators: this is usually acknowledged in terms of market share, as there is a huge premium for the first mover. But that inequality is well balanced by another: operators who can legally access larger data bases across different domains, and use them with less restrictions, will become more efficient than those which operate under antimonopoly, privacy or other constraints.

[12] Burrhus Frederic Skinner, *Beyond Freedom & Dignity* (Indianapolis, Ind.: Hackett Pub, 2002), cited by Shoshana Zuboff, *The Age of Surveillance Capitalism : The Fight for a Human Future at the New Frontier of Power* (New York: Publicaffairs, 2019), p. 439.

[13] Chris Anderson, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete," *Wired,* June 23, 2008, https://www.wired.com/2008/06/pb-theory/.

[14] Wolfie Christl, "Corporate Surveillance in Everyday Life - How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions," *Cracked Labs,* June 2017, https://crackedlabs.org/dl/CrackedLabs_Christl_CorporateSurveillance.pdf

Many other nations have a much more diverse starting point and agenda for the digital age than the United States. We will look at three cases: the European Union, and its top-down regulatory framework; China, with its statist vision of control and innovation coupled with dynamism from below; and India, which has characteristics from both the EU and China, while being very integrated within the American digital scene. Differences exist both from the technological, societal or regulatory points of view.

Beyond these cases, more general distinctions can be made: either countries are too small or undeveloped to regulate a digital space on which they have no hold. They are often market-oriented and they will go for efficient cost/benefit customer solutions, disregarding both privacy and government control over the content of communications. Or their governance is authoritarian, and they will choose digital packages that facilitate surveillance and eventual closure. A regulatory balance, as we shall see, which is the crux of the European choice so far, is not easy to define. But this is also very hard to implement, and it requires sophisticated rules and human resources.

This study identifies the key movers of the data privacy debate (I), studies its legal formulations (II), discusses our three cases studies (III, IV, and V), analyzes health as a theme-in-focus (VI), and concludes with outstanding issues (VII) and propositions for a data protection regime.

I

DEFINING THE ISSUE AND THE DEBATE

The goal of protecting personal data and privacy stands in a regulatory balance. This can be defined in very simple terms as a triangle, and as an Indian Supreme Court judge expressed in July 2018: "The citizen's rights have to be protected, the responsibilities of the states have to be defined but the data protection can't be at the cost of trade and industry."[15] But unpacking these easy definitions immediately compounds the difficulties. It is not only citizens' rights to privacy, but also company and IPR data that must be protected. The interest of the state is not only about national or public security, it also involves defining the public interest – including, as obvious examples, the benefits from health data banks vs. the patients' rights to privacy, the free media's right to investigate vs. the protection of the individual. Efficiency – or economic rationale – may imply increasingly large data banks across different domains, which not only challenge privacy but also create oligopolies.

Furthermore, this regulatory balance is a moving target. From the start of the digital age, technological or market innovation has tended to outpace rules. This is even truer of AI and other recent breakthroughs that rely on the computing performance of algorithms.

It is not surprising that the collection of data and the protection of privacy have become pervasive issues. It is a natural starting point, short of de-inventing the algorithm techniques that turn data into a

[15] ET Bureau, "Justice Srikrishna Committee Submits Report on Data Protection. Here're Its Top 10 Suggestions," *The Economic Times,* July 27, 2018, https://economictimes.indiatimes.com/articleshow/65164663.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.

weapon. Viewed with Shoshana Zuboff's accusatory words, "If new laws were to outlaw extraction operations, the surveillance model would implode."[16] Let's remember that the earlier legislation on data collection was a simple translation into the nascent computing age of reservations against filing individuals. Sweden's Data Act of 1973 required authorization from a national Data Authority for storing personal data, and then guidelines from this authority. France's text in 1978 – the Loi Informatique et Libertés, as it came to be called – was a one-pager banning the use for any public or private decision of data resting *on the sole use* of automated files profiling or containing information as to the personality of an individual. Judicial decisions could not use these automated files at all.[17] In short, the law neither banned the collection nor the use of files, but only their cross-referencing and derived decisions from them. These restrictions were original, yet they appear even more flimsy today than they were at the time.

But how far to go, and in which direction? What is certain is that data flows are too beneficial for global growth to be substantially hindered by regulation or control – unless a very deep cost is accepted. In the decade 2005-2014 only, data flows (information, searches, communications, transactions, video, and intracompany traffic) were multiplied by 45. Their contribution to the increase in GDP is greater than that from flows of goods. E-commerce also represented 12% of all traded goods in 2015. In fact, the advent of 3D manufacturing is likely *to reduce* trade in goods in the near future, as production will often be re-localized. This dry economic assessment does not include the multiple benefits and convenience

[16] Shoshana Zuboff, *The Age of Surveillance Capitalism : The Fight for a Human Future at the New Frontier of Power* (New York: Publicaffairs, 2019), p. 105.

[17] "Assemblée nationale et Sénat, "Loi N° 78-17 Du 6 Janvier 1978 Relative à l'informatique, Aux Fichiers et Aux Libertés", https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460

that the digital age will increasingly offer. Unlike labor or machines from the industrial age, information is scalable and is not exhausted when consumed. "The economics of digital information, in short, are the economics not of scarcity but of abundance. This is a fundamental shift, and a fundamentally beneficial one. If you have Internet access and a connected device today, it is both free and easy to keep in touch with the people who mean something to you—your kith and kin—even as you and they move around. As digital technologies make markets and businesses more efficient, they benefit all of us as consumers."[18] Even the asymmetry of information between operators described above is not complete. The diffusion of information technologies – if the required education needs have been met – can empower microfirms and companies generally considered to be outside the main production centers.

Awareness of their danger has come from experience with authoritarian states. George Orwell is the inescapable literary reference, along with its cinematographic homologue, *Black Mirror.* China's march towards a dystopia of gigantic data collection and use by the state is a case in point. But this reality is not confined to authoritarian systems. The laws may be different, but the tools are often very similar – if only because there has been much cross-breeding, for example, between Silicon Valley or the American digital scene, and its competitors which have appeared in China and sometimes dwarfed their models in size. To assess calmly what a "surveillance state" can achieve, one needs also to understand the tools already developed by what Shoshana Zuboff, for example, calls the new "surveillance capitalism."

[18] Erik Brynjolfsson and Andrew McAfee, *Race against the Machine: How the Revolution Is Accelerating Inovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy.* (Lexington, Mass.: Digital Frontier Press, 2011), p. 46.

Clearly, the legal environment makes a difference, even in a situation where data gathering cannot be de-invented, and where the available technologies create a huge asymmetry between the individual and the state – or the corporation. The use of Google's search engine can't put you in jail in our societies (leaving aside the "dark web"), while in Xinjiang (China), landing in the algorithms of the Integrated Joint Operations Platform (IJOP, 一体化联合作战平台 the phone app that collects data for the Ministry of Public Security) has meant internment in re-education camps for upwards of a million minority citizens.[19] Yet the tools – data collection, linking up various sources and using predictive algorithms – do have a distant relationship. It is even more obvious when one compares China's fast-rising social credit system with the decades old credit-rating system that is particularly prevalent in the United States. The last is constantly being refined by the introduction of new algorithms. Hundreds of thousands of Chinese are denied fast train, airline and other reservations because they have acquired a bad social credit rating which is unrelated to train travel, with little means to correct their score. In the United States, if you don't pay the insurance premiums on your car, an insurance company can turn off the engine from a distance and at the time of its choosing: this is bad enough, but at least the penalty is directly related to the financial default.[20]

[19] "China's Algorithms of Repression | Reverse Engineering a Xinjiang Police Mass Surveillance App," *Human Rights Watch,* May 1, 2019, https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance.

[20] Gary Hoffman, "Car Payment or Else: Engine Shut off Systems," *Autoblog,* December 19, 2016, https://www.autoblog.com/2009/06/27/engine-shut-off-systems/?guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAC8mFb3frqN6hIkaLhdvtwsLC0OW6P1A5Sn7mS8i9tvO6Lc8qsEm0fjWMPZk7SkW8_kirojwhauEJtoFnbnBbORFaKKrrjtqBaJnvN_0I2V_mNKAEniWBrGz1dOYbcTr5hZIHRDTODW-q6F2npEHCZeJSXExx6d9aiZqLTd8YLV5&guccounter=2

The U.S. Matrix

How societies deal with this challenge, involving both the positive uses of the digital age, and its downside and terrifying possibilities, is a question for everyone to consider. It has many angles, but that of privacy, and the debates surrounding this right, is a useful one because it sits at the intersection of the individual and the collective, of technology and regulation. It encapsulates large differences across societies and political systems, and in some ways, it resembles two earlier debates. In a narrow sense, it is those surrounding libel laws that arose with the print media. In a larger perspective, it is the debate for or against nuclear energy. Like the former, it is directly linked to the issue of freedom and its restrictions. Like the latter, it pits against one another the promoters of efficiency against those who prioritize more precautionary concerns. And it spans both debates in terms of scale. The digital age combines extreme centralization of data, platforms and operators, up to and including the Big State and giant companies, with an extremely wide dissemination of at least some of its features. Even before edge computing becomes commonplace, social media creates chain reactions that are akin to biological events such as epidemics. Rumors and fake news were the matter of earlier centuries, replaced by mass propaganda in the 20th century. They are back.

The privacy debate has two matrixes. One is clearly the United States. Digital technologies are largely invented there, the giant and not so giant companies that pioneer these have a global influence. Academic freedom together with a taste for relevant issues means that the very concept of privacy has largely been redefined there. The culture of litigation is equally important. Many test cases involving big data and privacy are arbitered by the courts, with precedent-setting

outcomes and large financial consequences. The prevailing wind has also shifted from blind faith in the promises of the digital age to awareness of its dangers for privacy and freedom. To some extent, the Apple generation (starting from a garage in 1976) directly followed the previous Woodstock Festival generation (1969), including its rebellion against prevailing corporate culture. Today, the next generation fights digital companies for privacy rights. The crusaders for digital privacy are also heirs to Ralph Nader and his life-long fight for consumers against corporations, from the Corvair.

America is therefore the mother of all privacy debates, and it has enacted important legislation in the past. Based on the Fourth Amendment ("each man's home is his castle" …), the Wiretap Statute (1968) and the Electronic Communications Privacy Act (ECPA, 1986) have defined the limits on collection and storage of data – with important loopholes and oversight, such as third-party use of personal data. Further legislation has often been about foreign intelligence and linked to, or justified by, terrorism, up until the revelations from the Snowden affair (2013). Yet, privacy has not been, at least during the last decades, at the forefront of federal legislation for several reasons. Instead there are sectoral regulatory regimes, especially in the health and financial data sector, which are patterned after the Federal Trade Commission's (FTC) fair trade principles. Privacy and its protection have largely remained for the courts to arbitrate – litigation and torts are a slow and procedural way to establish precedents from single cases, rather than adjudicate ex ante issues from above. This may not be less effective, as we shall see, but it is infinitely harder to describe, particularly when the judicial process is spread among 51 states. In fact, different rules bearing on digital privacy exist across these 51 states, without any of them being systematized. The federal government has been slow

to address, if not to recognize, fully the issue. Two successive administrations have each had their own priorities: under G.W. Bush, the fight against terrorism that coincides with a significant downgrading of privacy issues where national security is concerned. The Obama administration did not break with this trend. It was almost unanimously supported, and therefore influenced, by the Valley – that constellation of new business moguls, start-up entrepreneurs and talented geeks who have in common an unlimited faith in the digital age. In broad policy terms, it advocated for an approach to digital privacy rights that differs from the European approach, but it has not acted very decisively on its own recommendations.[21]

Both administrations have also recognized the extraordinary edge that new digital technologies give to the American economy – making up for what is lost in manufacturing. But the links between the Trump administration and Silicon Valley are far less symbiotic than was the case for its predecessor. It makes a new assessment of the situation. It may encourage antitrust decisions that hinder the horizontal growth of major platforms and digital companies across sectors – and that may place limits on big data aggregation, therefore favoring privacy rights. But the administration also lives in a context where competition with China is the first foreign priority. Moving with one arm tied behind one's back in big data, algorithms and the like is certainly not the preferred policy option. Again, interesting developments for privacy – including, as we shall see, the influence of the popular European General Data Protection Regulation (GDPR) – happen

[21] See especially, Executive Office of the President , "Big Data and Privacy: A Technological Perspective," *President's Council of Advisors on Science and Technology,* May 2014, https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf.

mostly at state level. At the level of Congress, the Cloud Act (2018)[22] is mostly an outcome of issues associated with international data-sharing. It is an example of the extra-territorial reach of American law, and constitutes a victory for the federal government over digital companies on the issue of turning over data stored outside the United States.

Inevitably, privacy debates therefore refer to companies, experts, and often judiciary decisions that originate in the U.S. Yet the American scene is so diverse that it defies any characterization, and hence is not the focus of our study. In one legal expert's view, "U.S. privacy law is a smorgasbord (…) sectoral statutes and torts cover narrowly defined behavior, and some additional constitutional proscriptions apply to government activity. But most private data-handling activities in the United States fall outside all these laws. In the absence of general-purpose omnibus privacy law like the E.U. Directive, consumer protection regulators such as the FTC and state attorneys general have moved in to fill the resulting vacuum."[23] The United States therefore remains a complex reference point.

There are significant differences between the American and European vision of privacy. "A consumer protection regime generally allows any collection and processing of personal data, unless it is specifically forbidden. Data protection law adopts the opposite default, permitting collection and processing only for a statutorily defined justification. In other words: in the United States, it is usually allowed unless the law specifically states that it isn't, while in the EU it is not allowed

[22] "Text - H.R.4943 - 115th Congress (2017-2018): CLOUD Act," Congress (2018), https://www.congress.gov/bill/115th-congress/house-bill/4943/text.

[23] William McGeveran, "Friending the Privacy Regulators," *Arizona Law Review* 58, no.4 (2016): 959–1026, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2820683.

unless the law says it is."[24] An echo of this is clearly visible in different views on default options.

Europe's Normative Power in Action

Europe has increasingly become the other major influencer on the debate. For a long time, Convention 108 was the only internationally binding agreement on data protection. It was opened for signature in January 1981 by the Council of Europe, and currently has 55 signatories, including non-member states. It forms a first basis for data protection, although in very general terms: Russia, for example, could sign it at this stage. The convention was modernized in 2018, with the aim of also harmonizing it with the GDPR. It now includes both manual and automatic processing, restricts states from creating exemption to processing under certain categories and updates breach notifications. It also creates a Convention Committee to monitor and supervise the application of the convention's principles by the parties. 33 states have signed the modernized Convention so far.

The European Union has reached another level of protection with its path-breaking GDPR, which came into force in May 2018. It is worth noting that in the 88 pages of the superbly written text, the word "privacy" occurs only twice – and in the same footnote, referring to a 2002 European Community directive on the protection of privacy in electronic communications (often dubbed the "cookie law"). The only other mention that comes close is that of "private and family life" that crops us in paragraph 4 of the Preamble. The regulation has a wider goal – protecting natural persons in the processing of

[24] *Ibid.*

their personal data and ensuring the free flow of data across the Union. Focusing on the collecting and processing of personal data (rather than on its use, an issue we shall come back to), it is a fine balancing act between the protection of individuals, the explicitly recognized commercial need for free flow of data in and out of the EU, and a series of exemptions where legal requirements (largely dealing with security) or the public interest (ranging from health research to the media's right to investigate) are concerned. Most of all, it is a top-down regulation that nonetheless takes in national exceptions. Unlike an EU directive, it supersedes existing national rules or laws (except where they might exceed the provisions of the GDPR) and mandates each member state to create a supervisory authority for implementation, processing complaints and reporting to the European Commission.

The GDPR is likely to be complemented by an e-Privacy regulation (replacing the 2002 directive), which is discussed since January 2017. This will have a broader impact than the previous directives: it will deal with the privacy and confidentiality of all electronic communications (including the new messaging and communications apps, for instance), and not only the commercial harvesting of personal data via cookies. Few people are aware that these new means of communication are prime sources for scraping personal data. As such, the e-Privacy regulation is still the object of much debate. It could impact much more severely the advertising business, while providing a more certain legal environment for telecom and messaging providers of what is called Over the Top (OTT) communications in terms of allowable metadata. The e-Privacy regulation is very slowly making its way through European institutions and will be debated with the newly elected EU Parliament.

To the individual digital user in Europe, GDPR is essentially a fairly unsystematic process of consent to cookies in order to access a website for the first time. Indeed, the first year is showing discrepancies across member states as well as issues regarding implementation. But this hardly renders justice to the obligations placed on "controllers" – the companies and institutions that store and manage personal data, for example. Furthermore, it is becoming an instance when the European normative approach has a wide influence on global debates, and influences rules in many other countries. This is not purely the result of moral suasion. The European market, and its digital flows, are huge. Accessing it is an economic necessity. Hence, to ensure the interest of member states, it is no surprise that even the new EU investment screening regulation, which entered into force in April 2019, listed "access to sensitive information, including personal data, or the ability to control such information" as determinants of a foreign direct investment's security or public risk.[25]

The EU has also put in place a regulation[26] that provides criteria for an "adequacy decision" regarding third countries or international organizations, thereby allowing the free transfer of data between the EU and those countries. So far, 13 countries have been recognized, with the United States being a special case as a "Privacy Shield" has been put in place that balances the impact of laws such as the

[25] "Regulation (EU) 2019/452 of the European Parliament and of the Council of 19 March 2019 Establishing a Framework for the Screening of Foreign Direct Investments into the Union," EUR-Lex, 2019, https://eur-lex.europa.eu/eli/reg/2019/452/oj.

[26] "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)," EUR-Lex, 2016, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2016.119.01.0001.01.ENG&toc=OJ%3AL%3A2016%3A119%3ATOC.

Cloud Act.[27] In preparation of its negotiation with the EU, Japan passed an Act on the Protection of Personal Information (APPI)[28] in 2017. Many more countries, including Brazil, India and South Korea, have applied. In America, some states – often those that are also taking the lead on environmental protection – are introducing similar legislation: the California Consumer Privacy Act (June 2018), the state of Washington Privacy Act (which failed to pass in April 2019). Texas, Massachusetts, New York and other states are considering similar bills. Leading academic experts on digital privacy routinely refer to the GDPR as an important precedent. One should add that many concepts and language used in the GDPR originate from American debates as well.

It is therefore undeniable that Europe is a key mover of privacy debates, and of the process of arbitrating between the three sides of the triangle described above – privacy, efficiency and security. As such, it will be the primary focus of our study. To supplement our understanding of data protection, we will simultaneously have two other cases, both countries which are increasingly becoming prominent in the digital debate.

India Between Models

The Indian case is the most clear-cut case for examining GDPR's international impact and potential as a model for recently proposed legislation on data protection. In using concepts like the control of

[27] "Adequacy Decisions," *European Commission,* 2019, https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en.

[28] Text and overview available at "Laws and Policies," Personal Information Protection Commission Japan, https://www.ppc.go.jp/en/legal/.

cross-border data transfers, notice and consent through privacy policies and the creation of a supervisory data protection authority, the Personal Data Protection Bill (henceforth, PDPB) appears to have been heavily modelled on the GDPR. India's digital sphere dwarfs that of Europe, in size if not in revenue. Like Europe, India's digital sphere is also dominated by outside actors – America's platforms, apps and software. In India's case, this is also shared with actors from China, who are already on top of the smartphone scene and tend to become very active in social media and gaming. The question of Indian firms gaining back the digital economy against powerful foreign companies is voiced more pressingly in India than in Europe. The similarities and the simultaneous differences between the European and Indian cases thus make India a significant comparison test for the European data protection regime.

Unlike Europe, India also has a constitutional system at the top, with a huge informal society and economy at the bottom. As noted by a draft national policy, "Today, two of three people in India do not have access to the kind of connectivity needed for digital trade and e-commerce. In addition, there is the problem of digital literacy and skills with only about 15% of rural households being digitally literate."[29] While it has continued to struggle with infrastructural impediments to digitalization, the number of internet users in India rose from 4 million in 2007 to 420 million in 2017 (equivalent to the users of the U.S., the U.K. and Germany combined). This is a constantly increasing consumer base, which is projected to reach

[29] Digital literacy is defined as "at least one person in the household who can use a computer, tablet or smartphone." Source: "Draft National E-Commerce Policy," February 16, 2019, p. 30, https://dipp.gov.in/sites/default/files/DraftNational_e-commerce_Policy_23February2019.pdf.

by 2025 the equivalent of the G7's combined users.[30] In addition to a large consumer base, India has important human resources in the digital area. In terms of industry and employment, digital India is much more closely intertwined with American companies than with Europe. In 2017, American companies generated more than 11 billion USD in revenue for India's huge data analytics industry; the UK generated 170 million USD (largest share in the EU) and the Netherlands produced 37 million (third largest share).[31] "Digital India" and related goals with proactive policies in this domain are a defining aspect of Narendra Modi's government. However, they have not been invented by this administration. Aadhaar, a unique national identification system derived from biometric and demographic data, was launched in 2010. An ambitious project, the Aadhaar card was also in many ways the start of the legal data privacy debate in India. The debate extends as well to the security lapses in the setting up of Aadhaar.[32]

While the proposed legislation also extends to private actors, there seems to be an undercurrent to the bill to protect local economy over foreign companies, especially American and Chinese tech companies that currently dominate the tech scene.[33] India seems to be looking towards the policy tools used by China in this regard. This is illustrated by the data localization requirement in the PDPB for example. "We don't want

[30] Rishi Iyengar, "The Future of the Internet Is Indian," *CNN,* November 27, 2018, https://edition.cnn.com/interactive/2018/11/business/internet-usage-india-future/.

[31] Source: Analytics India Magazine, AnalytixLabs, cited by Sandhya Keelery, "Infographic: India: Decoding Data for the Dollar," *Statista Infographics,* May 23, 2018, https://www.statista.com/chart/13935/decoding-data-for-the-dollar-india-analytics/.

[32] For an example: "Indian state government leaks thousands of Aadhaar names", *TechCrunch,* February 1, 2019, https://techcrunch.com/2019/01/31/aadhaar-data-leak/

[33] Newley Purnell, "India Looks to Curb U.S. Tech Giants' Power," *The Wall Street Journal,* August 13, 2018, https://www.wsj.com/articles/india-looks-to-curb-u-s-tech-giants-power-1534178721.

to build walls, but at the same time, we explicitly recognize and appreciate that data is a strategic asset," said Aruna Sundararajan, the nation's secretary of telecommunications, who has been deeply involved in the policy discussions for yet another proposed policy, the national e-commerce policy that calls for a "level playing field" for Indian companies.[34] The PDPB also offers exemptions for the state - often worded vaguely, offering broad, sweeping powers to the state machinery. Yet, India is a democratic, constitutional system and the executive has the judiciary to answer to.

In a way, India seems to be a bridge between the European and Chinese cases, modelling its legislation on the GDPR while using it as an instrument for its industrial policy. With its budding expertise and presence in the domain, it is also an important player to break the Sino-American hegemony in the digital sphere. Most importantly, in terms of adequacy, the Indian case remains significant for the EU. The developments in the latter's data protection regime could facilitate or disrupt trade flows between the two, as they heavily involve data transfers. The significance is also reflected in the European Commission's active participation in India's consultation process for data protection.[35] In subsequent parts, which are in no way a complete look at India's large and diverse digital scene, often ignored by Europeans, we attempt to view the legal situation and debates in India regarding privacy and the protection of personal data.

[34] Vindu Goel, "India Pushes Back Against Tech 'Colonization' by Internet Giants," *The New York Times,* August 31, 2018, https://www.nytimes.com/2018/08/31/technology/india-technology-american-giants.html?module=inline.

[35] Bruno Gencarelli, "Submission on Draft Personal Data Protection Bill of India 2018 by the Directorate-General for Justice & Consumers to the Ministry of Electronics and Information Technology (MeitY)," *European External Action Service,* September 29, 2018, https://eeas.europa.eu/delegations/india/53963/submission-draft-personal-data-protection-bill-india-2018-directorate-general-justice_en.

China's rules

The second case is China, previously alluded to in this chapter as an authoritarian dystopia in the digital world. It is a model to some in the technologies of surveillance, and in taming social media. China's quasi-intranet, considerably enhanced by recent rules governing international data transfer, is also a source of inspiration. "We must create a segment [of the internet] which depends on nobody", said Vladimir Putin, showing Russia's tendency towards a Chinese-style intranet.[36] On November 1, 2019, Russia's "sovereign internet" law has taken effect, giving the Russian government, in case of emergency, the power of switching its internet on and off from the outside world.[37] The Chinese rules are not meant to ensure data privacy, but to protect national security in the broadest sense possible, while keeping out undesirable information and opinion, also in the name of security. "Internet information service providers shall not produce, reproduce, distribute or disseminate information that (…) impairs national security, divulges State secrets, subverts State sovereignty or jeopardize national unity", stresses the 2000 Administrative Measures on Internet Information Service. The 2017 Cybersecurity Law clearly states in Article 1 that the law is formulated not only to ensure cyber security, but also to safeguard cyberspace sovereignty and national security. The 2018 Personal Information Security Specification exempts the need for consent when national security is concerned. The list goes on, and it is a visible feature of the Chinese rules that China does not try to hide.

[36] Andrew Roth, "Russia's Great Firewall: Is It Meant to Keep Information in – or Out?," *The Guardian,* April 28, 2019, https://www.theguardian.com/technology/2019/apr/28/russia-great-firewall-sovereign-internet-bill-keeping-information-in-or-out.

[37] "Russia Internet: Law Introducing New Controls Comes into Force," *BBC News,* November 1, 2019, https://www.bbc.com/news/world-europe-50259597.

China has vertical control over the internet and digital data. But uses are horizontal. Each of China's huge platforms cuts across different sectors, covering many different activities. This horizontal nature of China's companies hardly goes unnoticed. It is particularly evident with the stories of the social credit system widely discussed on Western media. Alibaba, the e-commerce company has successfully covered all aspects of the Chinese citizen through the service it provides, as well as acquiring shares in other companies. Through its own access and that of related companies, Alibaba has an unmatched view of Chinese customers across their life. The table below shows the main sectors covered.

Alibaba's Big Data: Internal Sources		
1	E-commerce data	Taobao, Tmall, Alibaba
2	Payment data	Alipay
3	Dating data	Wangwang, Laiwang
4	Video data	Youku
5	Browser data	Taobao Browser
6	Search data	etao
7	Game data	AliGame
8	Music data	Xiami Music
9	Travel data	Qyer
10	Map data	Gaode Map
11	ID data	Taobao Account

Alibaba's Big Data: External Sources		
	Purchased Party	Share of Alibaba
Search Engine	Yahoo China	40
Daily Life	Koubei	100
	Meiuan	10
	Kuaidi Dache	100
	Gaode Map	100
Social and Mobile Internet	Sina Weibo	18
	UC Browser	100
Culture	Xiami Music	100
	Culture China	60
	Wasu TV	20
	Youku Tudou	16
	Evergrande Football	50
	21st Century Media	20
Finance	Tianhong Asset Management	51
	Hundsun Technologies	100
Logistics	Singapore Post	10

Table Source: Chu Zhang and Leng Xin, "Research on the Application of Big Data in E-Commerce Enterprises 大数据在电商企业的应用研究," Journal of Chuzhou Vocational & Technical College 15, no. 5 (March 2018).

This is an advantage that non-Chinese companies find hard to compete with. The comparatively lower awareness of the Chinese regarding privacy and the willingness of the Chinese government to prioritize state objectives over any other consideration, make the Chinese case unique. This applies to national security with a broad and in fact unlimited definition. It is also exemplified by the state's emphasis on economic development, as shown by the case of the social credit system, which aims to create a trustworthy society that

facilitates growth, and by the "internet plus" strategy that aims to create new business sectors. The booming of the online economy has created new business opportunities: they include personal data theft and illegal sales. A recent *China Daily* article celebrates the detention of 7460 individuals by Guangdong police during a special campaign in the first eight months of 2019. As reported, 400 million items of stolen personal data, that was used by criminal gangs to defraud, were identified.[38] The case gives an indication of the magnitude of the issue regarding personal information protection.

The cases of China and India illustrate the tensions between privacy, marketing interest and state control, with very different solutions. Digital privacy is a priority that bumps against other goals, and Europeans must take into account that these goals are emphasized differently in China, India or the United States – the other three largest digital markets.

Innovation is one such goal, and in effect a requirement. The multiple applications and improved productivity that digital developments allow cannot be summed up adequately by the notions of efficiency or productivity: these are measures of economic performance, and the digital age delivers much more than just that. Furthermore, European regulations do not exist in a vacuum. They can hope to exercise leverage thanks to the size and attractiveness of the European data market, but innovation may thrive in less demanding regulatory environments.

[38] Caixiong Zheng, "Guangdong Police to Intensify Fight against Personal Data Theft - Chinadaily.Com.Cn," *Chinadaily,* September 19, 2019, https://www.chinadaily.com.cn/a/201909/19/WS5d833fd8a310cf3e3556c711.html.

Public interest is another requirement. Security or even public order do not encompass it completely – the example of the health sector provides an illustration, but we could have covered education, autonomous driving and many other promising areas. The requirements of public interest may clash with the objectives of data privacy: the issue is a two-way street. It is clear that the goal of ensuring privacy must find a compromise with that of delivering public goods, a definition that is wider than just public order or security.

II

WHAT IS PRIVACY AND HOW CAN IT BE ENSURED?

Data privacy is an umbrella term that is intuitively understood by everyone, but that is not easily defined. It is not confidentiality, nor is it cybersecurity, although it incorporates a bit of both. You may want to communicate information, including that of personal nature, to select people whilst not making that information public: it is the case with much personal data formerly stored on your thumbnail or hard drive, and now usually in a cloud. This is privacy combined with confidentiality. You may want that data to be safe from hacking by an individual or organization, and that is privacy associated with security. You may want your personal data to be treated anonymously (which is often the case with health information) in your immediate or long-term medical interest, but you do not want to see that data be used to assess, for instance, your insurance risk profile: this is about the use rather than the collection of personal data. Making decisions away from government eyes, and not suffering discrimination based on these decisions, or on your personal characteristics, is also closely tied with privacy.

Privacy is both an expanding bubble and a receding reality. In the 4th Amendment of the American Constitution, it is stated that “a man’s house is his castle”: much legal wrangling in recent decades has been about the right to search cars, and if these are extensions of “one man’s castle”, with varying interpretations. Guarding against intrusion is not the same as ensuring privacy however, and American law heavily balances privacy with freedom – including the freedom to investigate. In the European legal context, the very notion that the personal data of a human being can be traded is debatable: the

selling of our mind is no more acceptable to some than surrogate motherhood to many, or organs to most (the last is legal in Iran nevertheless, including for prisoners on death row: attitudes differ). In legal terms, privacy is expressed in terms of data protection. Personal data is the key focus of data protection regulations, and lies at the center of the privacy debate. According to the GDPR, personal data is "any information relating to an identified or identifiable natural person" either directly or indirectly.

Personal data can no more be separated from the individual than the mind from the body. But the opposition is in this case less prevalent because this is not a break-in or phishing. Opening mail, or even deciphering encrypted communications, are forms of intrusion, since the letter opener or code breaker are not legitimate receivers of these messages. Recombining information that has already become data, for instance with the help of algorithms, is an analytical rather than intrusive process. After one enters the domain of big data collection and treatment, an accepted distinction has long been made between data – whether digital or analog by source – that extracts information from individuals, and metadata that consists of descriptors of actual data: in the analogous world, recording a phone conversation and merely noting the time and duration of its occurrence between two individuals are not identical. For digital files, metadata allows storage, selection and communication of data with a guarantee of authenticity that is ensured by metadata standards.

What is one to make of data fusion from distinct data collection sources combined with algorithms? This combination transforms lead into gold, or isolated data points into recognizable identification and knowledge of individuals. There is no code breaking, no virtual

breaking and entering, and every step of the operation can be wholly legal, while the result is a trove of personal and often, sensitive data. These possibilities have been popularly illustrated for some time by the issue of anonymized data, or metadata. Data that has been anonymized can often – and increasingly, almost always – be de-anonymized in practice. In classic demonstrations, researchers were able to re-identify individuals by combining their zipcode, sex and birth date – with 84% certainty – or by matching movie searches on the IMDB website with Netflix data, or simply from released AOL search queries.[39] What is more, these examples show some of the ways in which the field of personally identifiable data has expanded – certainly much beyond an individual's ability to understand the consequences from the innocuous data that he surrenders as he leaves traces on the internet. Mood changes, but also the onset of Alzheimer's disease, can be inferred from keyboard click streams. In these cases, metadata morphs back into personal data, and the distinctions made by existing regulations can be voided. The alternative is to impose standards that are so wide that they create an obligation of results, and not simply of process, for the data operators. That, in fact, is often the case with the European GDPR that places the emphasis on results rather than on technical processes, but with ominous consequences: the prescription may not be implemented, or it may become a huge constraint on digital activities. It is, as we shall see, a political choice, the consequences of which have not yet been assessed and perhaps cannot be assessed fully.

[39] These three cases are well-documented in Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," *UCLA Law Review* 57 (2010), p. 1701, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

Still, the case for having regulation overshoot its target – legislating on results rather than on means – is straightforward. Current guarantees for privacy are risky. It is impossible to predict how data fusion, grouping together ever-expanding fields of data, and algorithms that have not yet been identified, may render obsolete today's technological or regulatory protection of privacy. In the words of a presidential commission under the Obama White House, "security deals with tomorrow's threats against today's platforms. That is hard enough. But privacy deals with tomorrow's threats against tomorrow's platforms."[40]

Anonymization and Pseudonymization of Personal Data

Methods towards de-identification have become more comprehensive, including for example, the obligation to use limited data sets to prevent data fusion on a large scale. Recent research proves this is becoming a harder task. It is likely that de-identification or anonymization techniques will have to become more and more sophisticated. Presently, with only five available data points and across 210 different populations, de-anonymization is achieved over 0.84 to 0.97 cases. With 15 data points, 99.98% of Americans can be re-identified. The results do not differ much if smaller samples are used. What's more, mean absolute failure is very low – under 0.041 when used on a 1% population sample. This is important because it means that "you know what you don't know." Failures to identify do not carry the same consequences as mistakes in

[40] Executive Office of the President , "Big Data and Privacy: A Technological Perspective," *President's Council of Advisors on Science and Technology,* May 2014, https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf.

identification. In other words, the results tend to be less and less deniable – reaching certainty as the number of data points grow.[41] A low estimate for digitalized populations is that each individual has created an average 5000 data points, a figure that will multiply with the uses permitted – and often not yet invented – for 5G and the IoT. Smart cities and smart homes will entail or technically allow the collection of infinitely more data points – think of video cameras as the most obvious example. One existing device, the Nest thermostat, learns what you like, "knows when you're away, learns about your home, controls it when you are away."[42] To achieve this, the device tracks movement, sound, electrical consumption of all devices. It only lacks front cameras – which other Nest devices employ.

From the above, one clearly sees that privacy is not ensured by anonymization alone. Yet, it includes this as an aspirational goal and as we shall see, anonymity features prominently in almost all privacy regulations. It is safer to adopt a very wide definition for privacy, an umbrella term covering "rules and norms on action and inaction related to our personal information."[43] Woodrow Hartzog further ties privacy to "trust, obscurity and autonomy." Trust is opposed to control – as we shall see, no individual can hope to manage and effectively control his/her personal data. Obscurity is preferable to secrecy – an unattainable goal again for individuals.

[41] Luc Rocher, Julien M. Hendrickx, and Yves-Alexandre de Montjoye, "Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models," *Nature Communications* 10, no. 1 (July 23, 2019), https://doi.org/10.1038/s41467-019-10933-3.

[42] Nest is now owned by Google. The claims are made at "Real Savings," *Nest,* https://nest.com/thermostats/real-savings/.

[43] Woodrow Hartzog, *Privacy's Blueprint : The Battle to Control the Design of New Technologies* (Cambridge: Harvard University Press, 2018), p. 10.

Autonomy is the preservation of a secluded garden for making up your own mind - including the right for individuals to engage in meaningful exchange among partners of choice.

Social Media and Personal Data

The digital age expands considerably our social horizon although, many will argue, that is at the expense of the thickness in relationships. One person's "meaningful relationships" max out at 150, according to a famous social psychologist, while our capacity for name recognition is said to extend to 2000.[44] Social networking sites use filter settings: Facebook is often said to limit that number at 5000, Twitter limits to 1000 the number of messages that you receive every day. This protects the ergonomy of the app. But algorithms also order the ranking of messages in your inbox, supposedly according to our observed preferences.

A form of privacy – data protection – is also a business requirement for data hoarders themselves. Facebook needs to protect its data simply because it sells their use. Conversely, the lure of free information and service over the internet is the biggest incentive to disregard one's privacy. Because internet is free of charge to users (except carriers' fees), it is paid by others: advertising, often dubbed "the original sin of the internet." From this flows a mirror conclusion: "if it is free, it means you are the product." In fact, some of the keywords for data extraction come straight out of the strip mining or

[44] Robin I. M. Dunbar, "The Social Brain Hypothesis," *Evolutionary Anthropology: Issues, News, and Reviews* 6, no. 5 (1998), p. 184, https://doi.org/10.1002/(sici)1520-6505(1998)6:5<178::aid-evan5>3.0.co;2-8. Cited by Woodrow Hartzog, *Privacy's Blueprint : The Battle to Control the Design of New Technologies* (Cambridge: Harvard University Press, 2018), p. 109.

meat packing industries: "data scraping", after which what's left of an individuality is the worthless "carcass." VPNs are a partial remedy hiding an individual user's real Internet Protocol (IP) (and not much more). There is a huge caveat – few users stop to contemplate who owns these VPNs and what secondary utility they might have. A 2019 study shows that a third of the world's top VPNs are Chinese owned, often through subsidiaries, Pakistan comes behind, with "the world's worst cyber-law", and VPNs based in the United States are obviously not free of surveillance of non-citizens.[45] In short, a VPN can ensure protection against some, but seldom against all.

In terms of government surveillance, one major default of intentional "backdoors" installed on hardware or software design is that if the backdoor exists, it can be also used by others – the proverbial "bad guys." This is akin to leaving the key under the doormat at your home. Governments themselves, quite outside of legal considerations, struggle under two contradictory requirements for security, whether it is in the area of encryption or that of its own covert access to data. For encryption, increasing the level of coding increases protection. Lengthy keys are also more cumbersome to use. The protection also works for illegal communications. France first required all encryption keys to be communicated to public authorities, then (from 1999 to 2004) only for keys above 128-bit. Similarly, installing backdoors (as may have been the case for Cisco routers as exposed by Edward Snowden) may create entry points for others; conversely, creating a foolproof architecture will lead to legal fights in order to access data: Apple gained a reputational advantage by denying the FBI access to the iPhone of a slain terrorist in the United States, at the expense of fighting terrorism.

[45] Jan Youngren, "Hidden VPN Owners Unveiled: 99 VPNs Run by 23 Companies | VPNpro," *VPNpro,* June 2, 2019, https://vpnpro.com/blog/hidden-vpn-owners-unveiled-97-vpns-23-companies/.

Localization of Data and Data Sovereignty

Since privacy is tied to collected data's security, one major issue has been the control over networks, and the location of, and access to, data clouds. The worldwide web – and in effect, digitalized data and communications – were built on the free flow of data across borders. The geographical location of servers – in the giant buildings that are in effect the cloud – is often based on opportunity costs (mostly the average local temperature and/or price of electricity) rather than on security factors. Figures vary greatly, but one account has 24 major companies slated to operate 420 data centers classified as hyperscale servers at the end of 2018, replacing not only your hard drive or localized company servers, but even the first generation of clouds.[46] Hyperscale centers allow more collection and linkage of data, and are likely to replace a lot of physical network gear, for instance in mobile communications. 45% of these centers were located in the United States in 2017, and 8% in China, its nearest competitor. The cloud solution leader, Microsoft, spends 15 billion USD each year on its cloud architecture, including the signature Azure brand. By way of comparison, the European Commission estimates that overall, 2 billion EUR in Horizon 2020 funding will be allocated to the European Cloud initiative over five years.[47] In France, two attempts at domestic clouds with public funding have failed.[48]

[46] Jeff Borker, "What Is Hyperscale?," *Digital Realty,* November 15, 2017, https://www.digitalrealty.com/blog/what-is-hyperscale.

[47] "The European Cloud Initiative," *European Commission*, August 17, 2018, https://ec.europa.eu/digital-single-market/en/%20european-cloud-initiative.

[48] Florian Dèbes, "Une Page Se Tourne Pour Le Cloud Souverain Français," *Les Echos,* August 1, 2019, https://www.lesechos.fr/tech-medias/hightech/une-page-se-tourne-pour-le-cloud-souverain-francais-1118112.

The issue of access and control over data has given way to government policies designed to localize data on their territory. Sovereignty over data is a politically sensitive theme. It is also one possible answer to another issue: data privatization by metaplatforms, which are increasingly taking over activities – health, education – that were functions of public systems. For better or for worse, states now delegate the set-up of their repetitive tasks – such as civil servant pay – to IT companies. Mapping systems are increasingly led by private actors such as Google and Apple. Only one nation, Estonia, has established the state as a multitask platform.[49] The issue of control over this privatized data looms large. China's solution is hybrid control ensured through largely customary means between theoretically private platforms and the Party-state, and rigorous control of all data transfer. Some other countries – such as India – mix acceptance of global platforms with at least an attempt to localize data inside the country.

Conversely, the U.S. Cloud Act (Clarifying Lawful Overseas Use of Data Act)[50] is a prime example of the extraterritorial reach of U.S. jurisdictions since it compels American companies storing data abroad to turn these over at the request of domestic law enforcement authorities. The requests cannot be in bulk. They have to be requested by court and on the basis of probable criminal cause (and not on the basis of national security for instance). Companies that claim conflict with a foreign law can refuse the data transfer – a very important provision because it essentially ties the implementation of the Cloud Act to the existence of compatible rules in other countries.

[49] See the 4-part series by Gilles Babinet, "The End of Nation States?," *Institut Montaigne,* 2019, https://www.institutmontaigne.org/en/series/end-nation-states.

[50] "Text - H.R.4943 - 115th Congress (2017-2018): CLOUD Act," Congress (2018), https://www.congress.gov/bill/115th-congress/house-bill/4943/text.

The Act in fact authorizes the executive branch to sign data exchange agreements with foreign governments. This has renewed a tussle on data flows across the Atlantic – which are still the largest data flows worldwide. A negotiated "Safe Harbour" treaty was earlier struck down by the Court of Justice of the European Union (CJEU). The U.S.-EU Privacy Shield Agreement, sometimes presented as a response to the Cloud Act, was in fact adopted earlier in 2016. It allows the free transfer of data – including personal data – to companies that are certified in the U.S. under the Privacy Shield, and is reviewed annually. It is also complemented by a U.S.-EU Data Protection Umbrella Agreement, essentially setting rules for the exchange of data between law enforcement authorities.[51]

Both the Cloud Act and the Privacy Shield are the subject of considerable polemics. The yearly Commission reports provide some perspective on outstanding issues: actual supervision of implementation by the U.S. Department of Commerce, the long-delayed nomination of an ombudsman that would provide permanent recourse for individuals and companies are currently the main sticky points. The Commission also "encourage(d) the U.S. to adopt a comprehensive system of privacy and data protection."[52]

[51] The Privacy Shield, the US-EU Umbrella Agreement and yearly reviews are helpfully available at "EU-US Data Transfers," *European Commission,* https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/eu-us-data-transfers_en.

[52] "The Second Annual Review of the Functioning of the EU-U.S. Privacy Shield," *European Commission,* December 19, 2018, https://ec.europa.eu/info/sites/info/files/report_on_the_second_annual_review_of_the_eu-us_privacy_shield_2018.pdf.

Privacy Policies, Notice and Consent

Notice-and-consent is one of the most widely used characterizations of consumer privacy, and obliges the collection and use of data be notified to consumer, in order for them to consent to it. This is done most commonly through privacy policies. Of course, the requirements for privacy can also diverge across societies and time. In a characteristically blunt way, Jack Ma, the Alibaba founder, professes to prefer Africa over Europe because "Europe is too concerned with privacy and rules."[53] For what they are worth – opinion surveys in China's controlled environment are a debatable enterprise, 38 % of the Chinese public would be willing to give up data privacy, usually in the interest of safety and trust in transactions. Aadhaar – India's wholesale filing of all citizens based in part on a biometric recognition system – would have met stronger opposition had it been applied in Europe (which nonetheless accepts more dispersed collection of personal data). A recent Australian study of internet platforms has collated Google privacy policies ab initio and tabulated the categories of personal data that Google is on the record for holding (which is not the same as publicizing, as the company emphasizes its in-house development and denies that it would sell identifiable data).

[53] Yunyu Qu, "Ma Yun: Europeans Worry Too Much, Alibaba Chooses Africa, Which Is More Willing to Believe in Technology 马云：欧洲人担忧太多，阿里选择更愿相信技术的非洲," *Caixin,* January 23, 2019, http://companies.caixin.com/2019-01-23/101373592.html.

Information Google disclosed in its Privacy Policy 1999-2019 as collected from users

	Jun 1999	Jul 2004	Jan 2009	Dec 2014	Jan 2019
Name	✘	✔	✔	✔	✔
Birthday	✘	✘	✘	✔	✔
Phone number	✘	✘	✘	✔	✔
Email adress	✘	✔	✔	✔	✔
Voice and audio information	✘	✘	✘	✘	✔
Payment information	✘	✔	✔	✔	✔
Location	✘	✘	✘	✔	✔
GPS	✘	✘	✘	✔	✔
Sensor data *via* wifi towers, bluetooth, etc	✘	✘	✘	✔	✔
IP addresss	✘	✔	✔	✔	✔
Your emails on Gmail (released Apr 2004)	NA	✘	✔	✔	✔
Your uploaded photos	✘	✘	✘	✔	✔
Your uploaded videos	✘	✘	✘	✔	✔
Your messages	✘	✘	✘	✔	✔
Your phone calls	✘	✘	✘	✔	✔
Comments you post	✘	✘	✘	✔	✔
Your calendar events on Google Calendar (general release Jul 2009)	NA	NA	NA	✔	✔
Your search history	✘	✘	✘	✔	✔
Videos you watch on Youtube (acquired Nov 2006)	NA	NA	✘	✔	✔
Devices you use	✘	✘	✘	✔	✔
Apps you installed	✘	✘	✘	✘	✔
Browsers you use	✘	✔	✔	✔	✔
This-party websites visited using Google's adversiting services	✘	✔	✔	✔	✔
Chrome browsing history (released Sep 2008)	NA	NA	✘	✔	✔

	Jun 1999	Jul 2004	Jan 2009	Dec 2014	Jan 2019
Browser information	✘	✔	✔	✔	✔
Device information	✘	✘	✘	✔	✔
Cookies generally	✔	✔	✔	✔	✔
Purchase activity	✘	✘	✘	✘	✔
DoubleClick cookie information (DoubleClick acquired Mar 2008)	NA	NA	✘	✔	✔
Mobile network information	✘	✘	✘	✔	✔

Source: Australian Competition and Consumer Commission, "Digital Platforms Inquiry Final Report," June 2019, p. 380, https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf.

One look at the above is enough to judge the merits of current privacy policies based on notification to individuals and consent. It would be counter-productive to abandon these obligations, reducing internet to a free- for- all hunt for your data. But the reality is that no one reads privacy policies, especially as it would be necessary to read hundreds of them, as well as their updates, and to understand the legalese behind them. The table also reveals how cookies, which were once the prime engine for data collection, are now just part of a much bigger picture. It is useful to select the case of the biggest worldwide platform: Google. But Google is far from unique – and third-party resellers of data from your clicks engage in far more egregious practices. With 7906 words, Google's latest statement of privacy policy[54] is also far from being the longest of its kind, although its text incorporates many clickable segments that open into new descriptions and more hard choices to make. What distinguishes a company such as Google is the size and breadth of its data reach. It is matched only by the biggest Chinese online companies. These

[54] As of January 22, 2019, "Privacy Policy," *Google,* https://policies.google.com/privacy?hl=en-US.

have far less restrictions on the types of businesses they can simultaneously engage in – including banking, insurance, and the burgeoning payment industry. Instead, the opportunity to acquire and consolidate personal data – or "behavioral surplus" – rather than size alone, is what drives many acquisitions by large American high-tech companies. What makes Google unique is the quality of its algorithms, and its financial ability to purchase other pioneering companies for their own algorithms and their domains of use. Two million companies depend on the marketing results that Google's big data and algorithms churn out. Yet, in many ways, it is more protective of the huge quantity of data that it acquires than many digital media and publishers. The latter rely directly on income from advertising to survive and therefore indiscriminately open their websites to third party brokers and resellers. Because of the need for online revenue, these news media outlets actually "stand up for a system of mass surveillance which goes as much against their readership as it goes against the journalistic profession", comments one privacy-oriented NGO.[55]

There is great interest from consumers for their data privacy. 90% of adults in the U.S. believe it is important to have control over what information is collected on them, 93% consider it important to be able to control who has access to this information,[56] and 86% have made efforts to hide their digital footprints;[57] these numbers resonate

[55] "EPrivacy Regulation: Do Not Let the EU Sell Our Right to Privacy," *La Quadrature du Net,* 2017, https://eprivacy.laquadrature.net/en/.

[56] Mary Madden and Lee Rainie, "Americans' Attitudes About Privacy, Security and Surveillance," *Pew Research Center,* May 20, 2015, https://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/.

[57] Lee Rainie et al., "Anonymity, Privacy, and Security Online," *Pew Research Center,* September 5, 2013, https://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/.

in Europe as per the 2018 Digital Attitudes report.[58] Yet, these opinion trends are inconsistent with actual online behavior. People often agree to share their data for free applications like Wi-Fi and websites. This is the privacy paradox. In many ways, it should be perceived as the privacy dilemma. Consumers are often left with the choice between surrendering their privacy and long privacy policies full of legalese.[59] One study suggested that an average American would have to spend more than 25 days a year to read all the privacy policies he or she was exposed to in a year.[60] Then, there is also the potential denial of goods or services if data collection is not consented to. Thus, the intuitive case that each individual knows best what to do with his or her own personal data is easily defeated. Scores of studies have now made this point – and yet, many regulations and the perception by the general public rest on the notions of notice and consent. As we shall see, efforts around the GDPR focus in part on improving the ergonomics of notice and consent, including standardization and universality of use: these efforts have merits. Not to underwrite them would lead to a worse situation for data privacy. Yet they only scratch the surface of the issue. In reality, it is impossible for a person to read the necessary notices, and even more impossible – *including for experts* – to comprehend what type of data – or metadata – is likely to be used in the future for intelligence about any individual.

[58] Joe Toscano, "Privacy By Design: What Changes Are Necessary, How To Do It, and How To Sell Your Boss," *Medium,* October 30, 2018, https://medium.com/greater-than-experience-design/privacy-by-design-7b1165d045e0.

[59] Kai Burkhardt, "The Privacy Paradox Is a Privacy Dilemma – Internet Citizen," *Internet Citizen,* August 24, 2018, https://blog.mozilla.org/internetcitizen/2018/08/24/the-privacy-paradox-is-a-privacy-dilemma/.

[60] Alexis C Madrigal, "Reading the Privacy Policies You Encounter in a Year Would Take 76 Work Days," *The Atlantic,* March 1, 2012, https://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/.

What is inaccessible to experts and to literate and experienced users is even less likely to be available to poorly educated users of smartphones, social media and other popular apps. Among open markets, India is set to become the largest smartphone and internet base, hotly contested by telcos and platforms. Semi-illiteracy encourages voice-driven apps, as seen in the increase of voice search queries on Google in India by 270% per year.[61] The likelihood that the general public can effectively manage a user-based privacy design based on consent and notice is close to zero.

Right to Be Forgotten

The same applies to other important components of privacy, such as the right to know what is known about you, or the "right to be forgotten", which are practical implementations of the need for obscurity to ensure privacy. Again, an ingenious piece of journalism about Google and Facebook highlights how unattainable these goals are.[62] For one person, the average downloaded amount of information from Google alone is 687,5 MB or 3 million words, or more than two volumes of the Encyclopedia Britannica. In practice, the answer to the question about what is known about you is simply "we really know a lot." Sorting through the data is even more difficult, as the same information is held redundantly by different sources. "From a policymaking perspective, the only viable assumption

[61] Rishi Iyengar, "The Future of the Internet Is Indian," *CNN,* November 27, 2018, https://edition.cnn.com/interactive/2018/11/business/internet-usage-india-future/.

[62] Dylan Curran, "Are You Ready? This Is All the Data Facebook and Google Have on You | Dylan Curran," *The Guardian,* December 19, 2018, https://www.theguardian.com/commentisfree/2018/mar/28/all-the-data-facebook-google-has-on-you-privacy.

today, and for the foreseeable future, is that data, once created, are permanent."[63]

Thus technology impairs the "right to be forgotten", one of the foundations of privacy law. In 2014, the CJEU ruled against Google, ordering it to remove from its list of results data relating to two Spanish citizens (an internet mention of a land sale that had happened years ago).[64] It is important that the Court mentioned that these were private citizens, and not public figures. This became the basis for the "right to be forgotten." Applying only to minors under 18, California's own "eraser law" entered into force in January 2015.[65]

A related development concerns the extent of data retention – and the public authorities' right to access them. Where companies see no value in data, they have no particular interest in retaining them, a costly process. The state's interest – justified by criminal cases, for example – may be much more extensive. In particular, metadata from phone calls, and phishing all communications from a cell tower in some cases, are important investigative tools. How much they can be used is a matter of debate. In December 2016, the CJEU ruled that retaining and accessing telco metadata (time of call, number called) is only appropriate in individual criminal investigation,

[63] Executive Office of the President, "Big Data and Privacy: A Technological Perspective," *President's Council of Advisors on Science and Technology,* May 2014, p. 40, https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf.

[64] C 131/12, Europa.eu (European Court of Justice 2014), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62012CJ0131

[65] "Senate Bill No. 568," California Legislative Information (2013), http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201320140SB568.

and cannot be performed in bulk.[66] The ruling is resisted by member states,[67] which have embarked on a slow consultation process to find acceptable alternatives. Finding a targeted method rather than blanket retention seems difficult to achieve.[68]

Even on a single platform or app, the only way to implement the right to be forgotten, enshrined by the so-called "eraser laws", is to push the nuclear button – delete ALL data, which will gain you a short moment of obscurity until you tap again a keyboard, open your phone, begin moving or start any other activity that is recorded. Even this short moment of obscurity is doubtful: the same data has likely been duplicated elsewhere – on your other machines, by partners, or of course by third parties beyond your reach. Again, Google and Facebook are the obvious guilty parties – some would say scapegoats, simply because they are so ubiquitous and so large in their range of monitored activities (although, once more, less so than their Chinese competitors). But a host of other platforms and apps, often unbeknownst to their users, would present similar issues with the additional challenge to simply identify them.

[66] This CJEU ruling on the Tele2 Sverige case is available at Court of Justice of the European Union, ECLI:EU:C:2016:970, *Europa.eu,* 2016, http://curia.europa.eu/juris/document/document.jsf?docid=186492&text=&dir=&doclang=EN&part=1&occ=first&mode=DOC&pageIndex=0&cid=4147224

[67] For a criticism of member state reluctance to accept the consequences of the CJEU ruling, see Jesper Lund, "EU Member States Willing to Retain Illegal Data Retention - EDRi," *European Digital Rights*, January 16, 2019, https://edri.org/eu-member-states-willing-to-retain-illegal-data-retention/.

[68] Working document submitted to the December 2018 Home and Justice Council, November 23, 2018, http://data.consilium.europa.eu/doc/document/ST-14319-2018-INIT/en/pdf

Privacy by Design

One must therefore move away from notice and consent, and from the ergonomics or user experience (UX) in these processes, to other methods ensuring data privacy. A generic concept is privacy by design, or "baking privacy" into the design of IT. This is actually an approach that is favored by large digital companies because it rests on technological prowess, which is more available to them. It does not deviate from the fundamental assertion that "code is law" (Lawrence Lessig), and from the belief that law should stay away from tech because law moves much more slowly and cannot catch up with innovation – or only by stifling or stopping it. Privacy by design gives maximum power to the choices made by software designers. An example is Apple's encryption of iPhones. These choices, like end-to-end encryption, create dilemmas for authorities that fear the ability of criminals to "go dark" – escape surveillance for instance. In this cat and mouse game however, the balance seems increasingly loaded by newly available technologies towards surveillance tools. Sicilian Mafia bosses once escaped detection by communicating solely from their hidden retreats with pizzi, scraps of paper. Today, they need to fear every piece of electronic equipment in their environment.

Some tools for ensuring data privacy also imply trade-offs with competition principles. Recently, several major data platforms and their CEOs have turned towards privacy laws, including both GDPR and the coming European e-privacy regulation. This may be part of an awakening process to the damage already done. But there is a market angle as well: throwing off the boat a host of third-party data handlers and resellers, widening the scope of platforms to new domains while ensuring the data stays inside the platform's black

box will indeed foster more privacy (assuming the company and its employees can be trusted), but it will also extend the power of algorithms based on even bigger and wider data banks. It is also perhaps how major Western platforms can hope to prevail in the challenge from Chinese platforms that may be entirely under the eye of their own government but are otherwise much less regulated, and are using this for international expansion. In turn, this has generated a counter-push by the U.S. Department of Justice, which is considering a break-up of Google and possibly other major IT platforms.

To prevent the asymmetry of knowledge among data operators, one popular suggestion has been to impose an open standard to the biggest platforms – they would be obliged to share categories of data with competitors: typically, Amazon, which is now able to design and market the best quality/price compromise for lightbulbs because it knows so well the preferences of its customers from their clicks and purchases, would have to share this marketing information with other lightbulb sellers and manufacturers. Attractive as the idea is to ensure symmetry of information, it does imply that large data resources will be turned over to any number of third parties. Enhancing competition in this case works against privacy by design.

Companies cannot achieve privacy by design if they don't know the rules and don't receive guidance. There will be no benchmarks for them and for users, no goalpost for litigation. One stark assessment contrasts a "choice of digital critical infrastructure suppliers who are muddling through security and privacy debates, or one who actively relegate those debates in favor of digital authoritarianism or Chinese interests."[69] For example, it notes that countries involved in Belt &

[69] "Is The Innovation Winter Coming? – Analysis," *Eurasia Review,* June 27, 2019, https://www.eurasiareview.com/27062019-is-the-innovation-winter-coming-analysis/.

Road connectivity projects may not have a choice, and that 18 states so far have chosen to use Chinese monitoring systems.

But legislating privacy by design is difficult, and perhaps antithetic to innovation itself. "Privacy by design is a lot of hype with very little substance. Although it has enormous potential to reset the imbalance between data collectors and users, it suffers from too much ambiguity."[70] The obligations of each actor must be clearly identified – lest, for example, different types of operators reject responsibilities on each other. If the law is expressed too broadly it becomes impractical; if it is defined too narrowly it will miss most of the target. Vague requirements and terms encourage superficial compliance and make it hard for users to know if the law is respected. It also leaves open the possibility that various institutions and individuals in charge of overseeing compliance have different, or even opposed, interpretations of the law. There must be a mixture of assigning responsibility to software designers and operators, as well as to organized users: if extremist groups abuse the potential of social media to turn it into a weapon, it is not the sole responsibility of the social media. There should be cooperative implementation of the law. Indeed, given the myriads of organizations involved in digital exchange, explaining the law and persuading should be practiced before *nudging* or punishing.

Only this iterative process between the law and the actors can prevent a major drawback of ex ante legislation: it will err by missing unseen privacy issues, or it will overshoot the target and inhibit digital activities by being too broad and vague out of perceived necessity. Europeans, one should add, are particularly vulnerable to the latter excess because of the popularity of the precautionary principle,

[70] Ari Ezra Waldman, "Privacy's Law of Design," *UC Irvine Law Review,* October 31, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3263000.

which is an anti-innovation social ethos. Excessive and unrealistic regulation generally leads to poor enforcement of the law. A major issue is also the choice between top-down law enforcement and a bottom-up approach, for example through individual complaints and appeal or the empowerment of civil society groups that can better represent the interests of individuals. This is particularly important for privacy concerns. A key objective of privacy by design and of its legal prescriptions should be to unburden the individual from choices he or she cannot make, either because they are undecipherable, or simply because there are too many of them. "Trust but verify": only the addition of enforcing institutions and representative civil society organizations can provide a permanent check on operators. This is also made more necessary by the constant changes in the digital world. Updates, patches, improvements, initially unplanned uses happen all the time, and they pose the same risks anew.

And yet the precautionary approach has its limits, as it is often based on current consensus and customs. Without radical technological innovations, our societies would remain frozen. There are very few enduring consensuses. Because the digital world is molding anew our social customs and personal habits, extending the reach of our minds as much as an exoskeleton expands our body's capacities, one should consider that pressing the "stop" button has ethical consequences that are as large as carrying on. In 2006, when Facebook had a largely student membership of 8 million, it introduced "News Feed" – a feature that allowed members to track their Facebook friends' activities in real time.[71] Hundreds of thousands of members

[71] Tracy Samantha Schmidt, "Inside the Backlash Against Facebook," TIME, September 6, 2006, http://content.time.com/time/nation/article/0,8599,1532225,00.html. Cited by William McGeveran, "Friending the Privacy Regulators," *Arizona Law Review* 58, no. 4 (2016), p. 1004, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2820683.

organized protests against this feature. Today, perhaps for the worst, News Feed is a prime attraction for the 2,4 billion Facebook users. It is very clear that the service rendered by ride-hailing services outweighs the obvious loss for privacy that is involved. Being evaluated and graded every time one takes a ride, is not something that would have been conceivable only a few years ago. It is now taken for granted, and increasingly so in other sectors of the gig and bartering economy. Libertarians can always ride the subway – provided they pay with cash rather than with a card, as many Hong Kong demonstrators chose to do in July 2019 in an effort to escape personal identification.

This libertarian market argument also has its limits. Users or consumers who "pay" with their data have next to zero knowledge of what will be done with their data and the risks may incur. The "market" is skewed by the denial of service that an operator can oppose, and by the absence of any "pricing" of the data surrendered.[72] We are back to the initial need to balance privacy with efficiency – and security. Compromises are needed, perhaps by all: one of the reasons users are turning away from traditional television to streaming services is because they are spared the constant advertising. But in return, they also surrender their privacy.

[72] Katherine J Strandburg, "Free Fall: The Online Market's Consumer Preference Disconnect," *University of Chicago Legal Forum* 2013, 5 (2013), https://chicagounbound.uchicago.edu/uclf/vol2013/iss1/5/.

III

GDPR, A EUROPEAN REGULATORY FEAT

The GDPR has become, in the space of a year, the most commonly used yardstick to gauge legal privacy protection – even though its scope is actually on operators and companies' handling of personal data.

There are reasons for this. It aims to create a one-stop shop for decisions on personal data protection (although directives and guidelines on implementation matter a great deal). 88 pages of superb writing starting with the goals and reach of the regulation (173 "recitals"), and proceeding to its actual provisions (99 articles). Being a regulation rather than the previous directive it replaces, it is law to the 28 member states, at least on "equivalent" terms, and it is therefore a means of achieving legal certainty in all member states of the EU, and over its external data flows. National rules can, under certain conditions exceed, but not underbid, the protections in place. The regulation addresses many if not all of the issues of privacy mentioned earlier in Part II. For instance, consent requirements have been strengthened compared with the previous European directive: it must be affirmative, clearly spelled out, reversible, and explicit for "sensitive" personal data. Different usage of personal data requires separate consent from the data subject. Personal data is defined broadly – including financial data and International Mobile Equipment Identity (IMEI) or IP addresses.

Data processing can only be performed under six specific legal requirements, including but not limited to consent. The right to erasure extends to cases where the holding of data is no longer necessary for its initial purpose. The right to data portability has

been introduced. The regulation addresses ergonomy, anonymization, responsibilities of data fiduciaries in general and operators (which are respectively described by GDPR as "controllers"[73] and "processors"),[74] a cooperative approach but also very stiff penalties, the inclusion of any entity dealing with the personal data of EU residents, as well as the thorny issue of equivalency with other data protection regimes are all addressed. Privacy by design has been incorporated into the GDPR: in effect, it befalls upon data controllers to enact "appropriate technical and organizational measures" to conform with the Regulation. The minimization of data collection and the limitation of access to those strictly needed for its processing are also obligations. Notification of data privacy breaches "where they result in a risk for the rights and freedoms" is mandatory within 72 hours. It is also mandatory for data controllers or processors to appoint data officers, who are responsible within the firm for education and compliance issues. Finally, the GDPR provides a practical framework for implementation – entrusted to national data institutions but with a new European Data Protection Board in charge of oversight. It has replaced the advisory Article 29 Working Party (WP29). This has become a key aspect of the GDPR.

The U.S. simply does not provide a model that can be replicated in any other society: its legislation is dispersed and conditioned on widely diverse state and federal laws and agencies, and with a large role for case-by-case litigation. The GDPR, including its scheme for equivalent national provisions and supervisory boards, is a Cartesian model by comparison. Inasmuch as it provides certainty with

[73] According to Article 4, controller is "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".

[74] According to Article 4, processor is "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller".

simplicity, it is easily transferable – at least for the principles. Finally, Europe is one of the four big digital data markets (along with China, India and the United States), and a forward role for the GDPR has important implications for the worldwide free flow of data. Whereas "equivalency" is the guide for implementation of the GDPR across member states, "adequacy" is the reference for concluding free data flow agreements with third parties. This is in recognition that the legal and societal environments differ, and that different processes may be employed to reach a level of protection that is adequate but not identical to EU standards. The adequacy decision is not transferable onwards to yet another country, and it is reached only after an iterative process between the EU and its partner, therefore leading to changes in the rules governing data protection in that country. The adequacy decision is subject to regular review and may be subject to subsequent improvements.[75]

The Obvious Downsides

There are downsides, of course. The GDPR is a catch-all text that is built on balancing opposite objectives. Indeed, it states that the protection of personal data "is not an absolute right." This right is bounded by "legitimate interest." The regulation objective is actually said to encourage the free flow of data – a claim that is disputed by those who see heavy obligations and litigation risks appearing on

[75] As an example, the adequacy decision concerning Japan is 48 pages and 28000 words long.

Source: European Commission, "Commission Implementing Decision (EU) 2019/419 of 23 January 2019 Pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the Adequate Protection of Personal Data by Japan under the Act on the Protection of Personal Information" (2019), EUR-Lex, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.076.01.0001.01.ENG&toc=OJ:L:2019:076:TOC.

the horizon. SMEs and organizations with less than 250 employees are spared the recording and book-keeping of their processing operations. This exemption is itself qualified by a firm's frequent recourse to processing, a useful exception to the exception. The former Cambridge Analytica, for example, numbered fewer than 250 employees.[76] A number of commendable rights for natural persons are nonetheless recognized: control by natural persons of their personal data, explicit and easy consent to data collection, minimal processing (e.g. when no other means are available), the right to rectification and to be forgotten, the portability of data, the principles of data protection by design and by default, and strict standards on human rights and rule of law for equivalency agreements with non-EU countries.

But the Regulation has almost no prescriptions for ergonomics – users' experience or UX. This is not uncommon for a legal text, yet it is clear that user experience research, guidelines for implementation and standard processes are necessary, just as the average user must understand road rules. It is under those conditions that the notion of giving natural persons control over their personal data can be at least partially envisaged. The Regulation has also literally taken no notice of AI: the words "algorithm" or "data fusion" do not appear. While required, anonymization, or pseudonymization, do not consider the new capacities of machine learning to defeat these processes.

[76] "Cambridge Analytica," *Crunchbase,* 2019, https://www.crunchbase.com/organization/cambridge-analytica#section-overview or https://www.owler.com/company/cambridgeanalytica.

The Exemptions – Public Interest and the Member States' Prerogatives

There is a large number of exclusions from the Regulation. The longest list appears in article 23, as follows:

"1. (…)(a) national security; (b) defence; (c) public security; 4.5.2016 L 119/46 Official Journal of the European Union EN (d) the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security; (e) other important objectives of general public interest of the Union or of a Member State, in particular an important economic or financial interest of the Union or of a Member State, including monetary, budgetary and taxation a matters, public health and social security; (f) the protection of judicial independence and judicial proceedings; (g) the prevention, investigation, detection and prosecution of breaches of ethics for regulated professions; (h) a monitoring, inspection or regulatory function connected, even occasionally, to the exercise of official authority in the cases referred to in points (a) to (e) and (g); (i) the protection of the data subject or the rights and freedoms of others; (j) the enforcement of civil law claims.

2. In particular, any legislative measure referred to in paragraph 1 shall contain specific provisions at least, where relevant, as to: (a) the purposes of the processing or categories of processing; (b) the categories of personal data; (c) the scope of the restrictions introduced; (d) the safeguards to prevent abuse or unlawful access or transfer; (e) the specification of the controller or categories of controllers; (f) the storage periods and the applicable safeguards

taking into account the nature, scope and purposes of the processing or categories of processing; (g) the risks to the rights and freedoms of data subjects; and (h) the right of data subjects to be informed about the restriction, unless that may be prejudicial to the purpose of the restriction."

The regulation also cannot be invoked against the archiving of data for historical, scientific and statistical purposes, an addition welcomed by researchers but one that puts in doubt, at least in an absolute sense, the "right to be forgotten." It is clear that the GDPR has been designed with private operators and data fiduciaries in mind more than governments and public authorities. In the above-mentioned triangle between privacy, efficiency and security, it has notably tipped the scale towards privacy by imposing obligations on operators and data fiduciaries – including public organizations. Yet, it has backed off from many decisions that could have lessened security or impaired public policy objectives in general. Once one looks at the GDPR through this perspective, the contrast with the more explicitly consumer-oriented rules or court decisions on privacy in the U.S. is less apparent. The regulation also makes important but only generally defined exceptions to data transfer rules outside the EU. The conclusion on this point is, as mentioned in Recital 114 and Articles 48-49, that the Commission has the upper hand in deciding case-by-case according to general principles.

Carrying a Big Stick – The Deterrent Effect

One area where the GDPR makes a significant difference is in the area of potential penalties. They are meant to be "effective, proportional and dissuasive." For the strict obligations accruing to the controller, processor and from the certification and monitoring bodies, they can reach 10 million EUR or 2% of their worldwide turnover, whichever is highest. On "basic principles", data subjects' rights, transfers outside the EU and non-compliance with a cessation order, it is 20 million EUR or 4% of turnover. This, as we shall see, got the attention of major firms in the first year of implementation, by setting the bar for penalties very high. Under Article 58, data processing across EU borders can also be precluded as a corrective measure.

Assessing GDPR, One Year After

After its first year in operation, the EU's GDPR is being reviewed and assessed by various sources – including the Commission and the European Data Protection Board (EDPB), as well as national supervisory authorities. It is in this context that the real impact of the GDPR can be judged. The text had many stated intentions and represented a balancing act. How it is implemented, even in the initial phase, tells us more about its impact. The role that it plays regarding data protection and privacy in other legislations and in shaping global attitudes, including in the U.S., can also be assessed. By contrast, the relevance of China's laws and rules for privacy, with generally different stated objectives, seems to be only a sideshow in what is now the world's most advanced overt surveillance state. Also, the distance between legal texts and practices is such that it is the latter that matters.

Feedback and National Siloes

As should be the case for a regulation ensuring personal data protection, the EDPB's main yardstick to measure the GDPR's success after a year has been the awareness of citizens and their readiness to initiate complaints. In March 2019, 67% of polled EU citizens were aware of the GDPR's existence, and 57% were aware of national supervisory authorities, a commendable result. 281,088 queries to national data boards in the same time, of which 144,376 complaints and 89,271 data breach reports, sounds impressive.

One major innovation of the GDPR is to provide for cooperation among national boards for cross-border cases. Cases go inside an existing Internal Market Information system (IMI). Beyond mutual assistance (444 requests were made under that heading), many of these cases go through a one-stop shop mechanism, where a Lead Supervisory Authority must first be designated. So far, the actual figures are less impressive here. As of March 2019, there were a mere 466 reported cross-border cases, of which only 19 had found a solution. Of the total, 45 were one-stop shop cases, of which 6 reached a final decision. The EDPB also gives consistency opinions to ensure across the board equivalency among supervisory authorities. 29 such opinions have been given. The one-stop shop does not apply in cases where the entity in question operates from outside the EU: it is then liable in front of every national authority.

Implicitly, the EDPB and the Commission recognize some of the complaints of stakeholders. One is on the practical differences that persist among member states in ensuring adequacy to the Regulation. After one year, three member states have still not changed their legislation. The Commission keeps their names quite close to its

chest, but they are Greece, Portugal and Slovenia. Differences such as the minimum age of children for consent create difficulties for transnational platforms. Although the Commission currently emphasizes dialogue over sanctions, it does not always recognize the lack of clarity and practicality from some of the supervisory authorities. In Poland, PUODO, the supervisory authority fined a Swedish platform (Bisnode) 220,000 EUR (in fact the third largest fine based on the GDPR in the first year) for having failed to contact, by registered mail, six million people about their personal data acquired from public registers. The legal fight pits strong defenders of the GDPR's Article 14 against those who would argue for proportionality: the cost of six million registered letters would be prohibitive.[77] Neither PUODO nor the GDPR itself pay much attention to the practicality of the rules.

Where is That Big Stick?

Overall, 56 million EUR have been imposed in fines during the first year, including 50 million for a single fine by CNIL, the French national board, on Google (but that was appealed). The United Kingdom has gone further in August 2019, with a 99 million GBP fine against Marriott and a 183 million GBP against British Airways: so far, the departing U.K. is the strictest enforcer of GDPR sanctions! These figures tell their own story. Although apparently high, the number of complaints has a very low ratio compared to overall usage. Implementation remains largely within borders. The 4% of global turnover magic weapon has turned into very limited fines, sometimes

[77] Karolina Gałezowska, "Why You Should Pay Close Attention to the Polish DPA's First GDPR Fine," *Iapp.org,* April 22, 2019, https://iapp.org/news/a/polish-dpas-first-fine-pay-close-attention/.

described as fit for a first year of implementation. Much more encouragingly, companies have spent and hired to be in compliance with the GDPR. The International Association of Privacy Professionals (IAPP) estimates that after one year, 500,000 organizations have registered data protection officers in Europe.[78] Large IT companies in fact emphasize their investment to comply with the Regulation – Microsoft for example claims to have used 1600 engineer man-years to establish compliance and implement the GDPR rules world-wide, and not only in relation to Europe.

The overall number of complaints, especially cross-border complaints, the slow rate of resolution and the minimal amount of fines issued does raise a question of implementation. The EDPB notes that national supervisory authorities hardly received the budget increases that would be commensurate with their new tasks. Five of them have actually seen a decrease (Poland and the Czech Republic) or no increase (Austria, Belgium and Latvia). For eight supervisory authorities, the number of personnel did not change, for one (Czech Republic) it actually decreased.[79]

[78] For an explanation of the estimate: IAPP, "Approaching One Year GDPR Anniversary, IAPP Reports Estimated 500,000 Organizations Registered DPOs in Europe," *Iapp.org* (May 16, 2019), https://iapp.org/about/approaching-one-year-gdpr-anniversary-iapp-reports-estimated-500000-organizations-registered-dpos-in-europe/.

[79] For actual numbers, see: EDPB, "First Overview on the Implementation of the GDPR and the Roles and Means of the National Supervisory Authorities," *European Data Protection Board,* March 8, 2019, p.11-12, https://edpb.europa.eu/sites/edpb/files/files/file1/19_2019_edpb_written_report_to_libe_en.pdf.

The View From Companies and the Issue of Size

The overall attitude of large companies towards the GDPR is at least outwardly positive – although one major sharing platform interviewed has explained that no major company would come out publicly against the GDPR as a whole, but emphasized that user satisfaction and their security were the company's priorities over privacy rules. Complaints focus on the varying "equivalent" legislations of each member state, with little unification in sight. Time is therefore lost, and eventually fines incurred because of these differences. Not all data processors and operators are in equal position. It is clear that third party data brokers, and paradoxically smaller websites and apps (including the most reputable news media that have moved to free content online editions) are the most vulnerable to restrictions from notice and consent, since their revenue stream is strictly founded on reselling data turned over by users. One particularly egregious approach of the GDPR's "notice and consent" rules is the so-called Oath Privacy platform, to which a number of respected publications such as the Huffington Post belong, and which was originally owned by Yahoo and AOL. It now has 43 "foundational partners" scraping your data. To deal with privacy controls, one needs to read and click through a maze of screens, including a list of privacy policies arranged by 45 different countries, to then pore through the fine print of each cookie owner, finally moving on to a privacy dashboard.[80] In short, Oath it has made it practically impossible to implement the rules of the GDPR on its associated websites.

[80] As experimented on Huffington Post website on August 19, 2019/ First Oath pop-up screen at: https://consent.yahoo.com/collectConsent?sessionId=3_cc-session_0bceacc7-b5ad-48f1-ae44-bfde36ab129f&lang=en-us&inline=false

An interesting study using machine learning to test the declared privacy policies of fourteen major web companies (Google, Facebook, Instagram, Amazon, Apple, Microsoft, WhatsApp, Twitter, Uber, AirBnB, Booking.com, Skyscanner, Netflix, Steam and Epic Games) has found them all in default with the GDPR. "The evaluated corpus, comprising 3658 sentences (80.398 words) contains 401 sentences (11.0%) which we marked as containing unclear language, and 1240 sentences (33.9%) that we marked as potentially unlawful clause, i.e. either a 'problematic processing' clause, or an 'insufficient information' clause (under Articles 13 and 14 of the GDPR)", according to the authors of this report,[81] which is also cited by the European Commission's one year after multistakeholder report.[82]

These remarks should be balanced with the perceived clash between clarity and inclusiveness, given the complex requirements of the GDPR. In the Commission's multistakeholders' report, the remarks from companies and civil society or consumer organizations often point in opposite directions. Reading through the lines of this polite but nonetheless candid account, one could well be a spectator of an imaginary tennis match between two opposite teams. Perhaps surprisingly, the corporate stakeholders, while acknowledging the work necessitated by the GDPR, also recognize positive results – particularly for risk-based assessments of digital data processing.

[81] Giuseppe Contissa et al., "CLAUDETTE Meets GDPR. Automating the Evaluation of Privacy Policies Using Artificial Intelligence" ed. CLAUDETTE(https://www.beuc.eu/publications/beuc-x-2018-066_claudette_meets_gdpr_report.pdf, July 2, 2018)

[82] Multistakeholder Expert Group to support the application of Regulation (EU) 2016/679, "Contribution from the Multistakeholder Expert Group to the Stock-Taking Exercise of June 2019 on One Year of GDPR Application," ed. European Commission (https://ec.europa.eu/commission/sites/beta-political/files/report_from_multistakeholder_expert_group_on_gdpr_application.pdf, June 13, 2019)

SMEs and public sector stakeholders, however, remain more weary by the efforts involved.[83]

Yet, the GDPR has had immediate measurable effects on third-party data collection by cookies on websites. A study of the first three months of implementation across seven EU countries reveals a 22% drop on news websites, with the largest reduction in the United Kingdom (45 %) and the smallest in Germany (6 %). Among these, the large platforms – Google, Amazon, Facebook and Twitter – had very little reduction of presence.[84]

Integrated platforms – such as Google and Facebook – brew their own stew with algorithms and sell the product of the analysis, not the raw data; they are in a sufficiently commanding position with their users so that they do not run a risk of being denied much personal data. This can be a huge business. With a tiny office in Shenzhen, Facebook, which cannot be accessed from China, nonetheless gathers 5 billion USD in advertising revenue (9 % of its yearly turnover) from Chinese advertisers to Facebook's international users.[85] For major e-commerce platforms such as Amazon, the risk is less that of infringements on privacy than of an unequal competition: Amazon acquires – and utilizes – more data on clients' tastes than any single manufacturer could hope to have.

[83] *Ibid.,* p. 14-15.

[84] Timothy Libert, Lucas Graves, and Rasmus Kleis Nielsen, "Changes in Third-Party Content on European News Websites after GDPR," ed. Reuters Institute for the Study of Journalism (https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-08/Changes%20in%20Third-Party%20Content%20on%20European%20News%20Websites%20after%20GDPR_0_0.pdf, August 2018).

[85] The 9 % estimate is based on Facebook's latest turnover figures.

Source: Paul Mozur and Lin Qiqing, "How Facebook's Tiny China Sales Floor Helps Generate Big Ad Money," *The New York Times,* February 7, 2019, https://www.nytimes.com/2019/02/07/technology/facebook-china-internet.html.

Data network managers, whether telecom operators such as Orange or infrastructure managers such as Microsoft draw their revenues from the direct service they provide, including data security, and are therefore less sensitive to privacy rules.

One issue that large companies acknowledge without much hesitation, however, is that they are actually favored over smaller firms by the GDPR requirements, simply because they have more financial means and more human resources to deal with these requirements. They can also maneuver more easily. The priority that the GDPR gives to notice and consent leaves them room. Google has shifted the onus of asking for consent to its external data suppliers. One interviewed digital infrastructure company makes a distinction between its large and smaller clients: it is easy to partner with the former for the GDPR but much harder with the latter. This was anticipated by the GDPR's makers, who impose less requirements on companies under 250 employees. A recent Commission report blames a national German data supervision authority for unilaterally lowering that ceiling to 20 employees.

Learning to Love GDPR – And Still Hating the E-Privacy Directive That Comes Next

The turnaround by large IT companies is spectacular. Mark Zuckerberg has declared himself in favor of the GDPR.[86] So has Sundai Pichai, Google's CEO,[87] or Satya Nadella, Microsoft's CEO, who calls the

[86] Henry Farrell, "Facebook Is Finally Learning to Love Privacy Laws," *Financial Times,* April 4, 2019, https://www.ft.com/content/67b25894-5621-11e9-8b71-f5b0066105fe.

[87] Jon Porter, "Google's Sundar Pichai Snipes at Apple with Privacy Defense," *The Verge*, May 8, 2019, https://www.theverge.com/2019/5/8/18536604/google-sundar-pichai-privacy-op-ed-nyt-regulation-apple-cook-advertising-targeting-user-data.

GDPR "a fantastic start"[88] and advocates for a worldwide standard. Apple's Tim Cook goes several steps ahead in endorsing a federal privacy law, including the issue of AI, which, as we have seen, is almost entirely missing from the GDPR. He also distances himself from much of the industry, pointing out Apple's "healthy suspicion of authority" which is part of its consumer appeal: "Some oppose any form of privacy legislation. Others will endorse reform in public, and then resist and undermine it behind closed doors."[89] The barb is not without some reason. In the current climate where issues from the Edward Snowden affair to fake news and Cambridge Analytica have turned opinion, the GDPR may appear as a reasonably certain regulation with a limited scope. In particular, it has not much to say about the scale of data controllers, which can greatly limit competition. It is also largely focused on the user's rights – which the best equipped companies know how to circumvent. In short, what was a bogeyman on the horizon in 2015 is now a limited attempt at ensuring privacy. Indeed, as the IT world predicted, "technology is law" and it is moving on ruthlessly. The current concern of those companies that trade data is more with the European Commission's next step – an e-privacy regulation that will renovate the regulation of telecom data with a much larger scope than the former 2002 directive on "private life and telecommunications.". Because the regulation would require end-user consent for transmissions including

[88] Isobel Asher Hamilton, "Microsoft CEO Satya Nadella Made a Global Call for Countries to Come Together to Create New GDPR-Style Data Privacy Laws," *Business Insider France,* January 24, 2019, http://www.businessinsider.fr/us/satya-nadella-on-gdpr-2019-1.

[89] The transcript for Tim Cook's speech at the 2018 International Conference of Data Protection and Privacy Commissioners in Brussels on October 24, 2018 is available at: Jonny Evans, "Complete Transcript, Video of Apple CEO Tim Cook's EU Privacy Speech," *Computerworld,* October 24, 2018, https://www.computerworld.com/article/3315623/complete-transcript-video-of-apple-ceo-tim-cooks-eu-privacy-speech.html.

the IoT (potentially hundreds of devices per user), it goes beyond the "cookie law" it replaces and is now the object of much lobbying. In effect, that piece of legislation has been stalled at the Council since it was proposed by the Commission in October 2017.[90]

The Pitfalls of User Consent

The gap between large and small companies or organizations is especially visible in terms of the users' experience of notice and consent implementation. Larger platforms and entities, which draw users on multiple occasions and with different activities, tend to have an initial approval system at the entry gate to their services, based on an often overflowing privacy notice. It will then be regularly updated, usually with similarly large new notices whose content may differ only on a few – but perhaps critical – points. The vast majority of users will neglect reading these huge privacy notices, partly as a consequence of trust in a brand name, partly because of time constraints. And they will hardly, if at all, read the frequent updates. Smaller organizations that draw varied visits, often on a single or quasi-single basis, usually implement a questionnaire-based approach. The design of these varies greatly, with many infringing the spirit if not the law of the GDPR. Having to go through a list involving dozens, if not hundreds of third-party partners, and ticking them off one by one, is clearly a formidable task, especially if it is repeated on other visits and on a large number of websites. Others simply give the option of reading long privacy notices or explanations

[90] European Parliament and Council, "Proposal for a Regulation of the European Parliament and the Council Concerning the Respect for Private Life and the Protection of Personal Data in Electronic Communications and Repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)" (2017), EUR-Lex, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017PC0010&from=FR.

of their policies, but the only choice available is to accept it or leave the website.

The GDPR itself, requiring separate consent for different uses of personal data, without prescribing a single framework for questions and answers, has involuntarily contributed to this complexity. Data processors will then use so-called "dark patterns" and nudging in order to encourage users to give away consent. In fact, every one of us has experienced huge differences across websites – from those that provide a setting by default ensuring privacy and asking merely to accept this, to those requiring item by item changes, or those that require you to review long-winded privacy policies – perhaps on different websites – and finally to those who provide you with a take-it- or- leave -it option, effectively blackmailing visitors.[91] Because of the "privacy paradox" (users will willingly indulge, for reasons of efficiency, in practices that endanger their privacy), the GDPR is reaching only one part of its stated goals, even in the area of notice and consent. The Commission's recent victory communiqué on the GDPR[92] does recognize that 44% of users haven't changed their privacy settings since the GDPR was introduced. In fact, just as there are rules for driving on the road, there should be publicized rules for the virtual road and convenient, easy-to-implement processes. Ticking boxes is only the beginning of what should lead to a mutually

[91] For a scathing look at several major platforms' implementation of GDPR in the first months after it was put in force, see: Forbrukerrådet, "Deceived by Design. How Tech Companies Use Dark Patterns to Discourage Us from Exercising Our Rights to Privacy," Oslo, June 27, 2018, https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf.

[92] European Commission, "Communication from the Commission to the European Parliament and the Council. Data Protection Rules as a Trust-Enabler in the EU and Beyond - Taking Stock," July 24, 2019, https://ec.europa.eu/commission/sites/beta-political/files/communication_from_the_commission_to_the_european_parliament_and_the_council.pdf.

accepted privacy relationship between providers and final users. It comes back to the notion that the user cannot be saddled with the responsibility of protecting its own privacy.

The GDPR and the Possibility of Global Convergence

As we leave Europe, and especially given the huge markets in countries where the average literacy is lower, the need for privacy protection becomes even more apparent. One of the early successes of the GDPR is the number of legislations that have recently been based on parts and concepts of the GDPR. Another is the interest of other states in an adequacy decision from the EU, allowing for the free flow of data transfer to the country involved. With some hype, the Commission concludes that this led to "a global convergence of data protection rules (…) These laws often have a number of common features that are shared by the EU data protection regime, such as an overarching legislation rather than sector by sector rules, enforceable individual rights and an independent supervisory authority. This trend is truly global, running from South Korea to Brazil, from Chile to Thailand, from India to Indonesia."[93] The Commission recognizes this is not a one-size-fit-all issue, and cites other potential models such as Japan's "Data Free Flow with Trust" (DFFT) initiative, launched by Shinzo Abe at the Osaka G20 summit in June 2019. Japan's project, however, does not cover the transfer of personal data.

[93] *Ibid.,* p.1 and p.11.

The Commission's report cites adequacy decisions such as the EU-Japan Agreement as the avenue with the greatest potential.[94] It touches briefly and indirectly on the excessive reach by other states – mentioning negotiations about the Passenger Name Registration (PNR) process, and the issue of sharing electronic evidence in criminal investigations: data sharing with third countries for law enforcement falls under the EU's separate Police Directive.[95] So far, the EU has made 13 adequacy decisions, including two limited agreements with the United States and Canada. Out of the other 11, 6 are financial centers or even off-shore markets such as Andorra or the Faroe islands. The Commission is officially negotiating with South Korea. Other countries, such as India, Brazil, Indonesia, have expressed interest in an adequacy decision from the EU. More countries are adopting GDPR-like legislation in principle. Without a human framework and resources for implementation, this can remain cosmetic, even if it testifies to the general appeal of the legislation.

[94] For a description of the process leading to the EU-Japan adequacy agreement, see: Hiroshi Miyashita, "The Impact of GDPR in Japan," in *National Adaptations of the GDPR* (Luxemburg: Collection Open Access Book, Blogdroiteuropeen, 2019), 122–27, https://blogdroiteuropeen.files.wordpress.com/2019/02/national-adaptations-of-the-gdpr-final-version-27-february-1.pdf.

[95] European Parliament and Council, "Directive (EU) 2016/680 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 by Competent Authorities for the Purposes of the Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and on the Free Movement of Such Data, and Repealing Council Framework Decision 2008/977/JHA" (2016), https://publications.europa.eu/en/publication-detail/-/publication/182703d1-11bd-11e6-ba9a-01aa75ed71a1/language-en.

IV

INDIA, A DIGITAL BLEND

The Constitutional Right to Privacy, a Contested Issue

The debates around data privacy in India first came to light with the debates around the Aadhaar ID card. The biometric database has been broken into on multiple occasions. It also became compulsory in order to access certain public services and benefits, and personal data has been made accessible to private companies as well.[96] Against this background, a landmark decision by India's Supreme Court has shaped the privacy debate. In 2017, a nine-judge panel pronounced a decision that recognized privacy as a constitutional right, and directed the government to create a special committee facilitating the creation of a data protection regime in India. The Srikrishna Committee was formed to draft the Personal Data Protection Bill (PDPB) in 2018. The Bill has not yet passed the legislative stage, at which point it will become a legally binding Act.

The Srikrishna Committee also published a report the same year that turned the issue around, putting privacy in the context of a "free and fair digital economy" and "empowering Indians", as its motivations were clearly reflected in its very title. The draft bill itself lowers the bar on interpreting the Supreme Court's ruling, stating (as the GDPR in fact does) that the right to privacy is not absolute. This is not a complete surprise. According to one lawyer representing the plaintiffs

[96] Bloomberg, "Amazon's Real Rival in India Isn't Walmart," *The Economic Times,* August 16, 2018, https://economictimes.indiatimes.com/industry/services/retail/amazons-real-rival-in-india-isnt-walmart/articleshow/65418425.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.

in the original appeal to the Supreme Court, after one year "the judgement has done very little in terms of altering state practice."[97] The report sheds light on the motivations for the PDPB, and posits the U.S., the EU and China as the three possible paths: The U.S. having a laissez-faire system, the EU a consumer protection regulation approach, and China as emphasizing data protection as a means to ensure national security. From these choices, it goes on to propose a "synthetic fourth path", stressing that the proposed bill "protects individual privacy, ensures autonomy, allows data flows for a growing data ecosystem and creates a free and fair digital economy."[98]

The Proposed Bill, À la GDPR

The proposed PDPB has an extensive reach, being applicable to data collected or processed not only within the Indian territory by both Indian and foreign data fiduciaries, but also outside of India if it pertains to Indian citizens. There is also a last clause, which is vague but goes beyond the scope of the GDPR in dealing with data collected or processed in connection with any business that is carried outside of India. It heavily follows the GDPR model, laying down obligations for data fiduciaries and data processors,[99] while outlining the rights of individuals (data principals). These rights include confirmation and access, correction, data portability, as well as the

[97] Apoorva Mandhani, "The Right To Privacy Judgment Is A Year Old, But Not A Year Wiser," *Livelaw.In,* August 24, 2018, www.livelaw.in/the-right-to-privacy-judgment-is-a-year-old-but-not-a-year-wiser/.

[98] Committee of Experts under the Chairmanship of Justice B.N. Srikrishna, "A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians," (July 27, 2018), https://meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf.

[99] According to the PDPB, "Data processor" means any person, including the State, a company, any juristic entity or any individual who processes personal data on behalf of a data fiduciary, but does not include an employee of the data fiduciary.

right to data erasure. The data fiduciaries include government bodies and related public entities, which is a step further from the pre-existing data protection obligations under Information Technology (Amendment) Act, 2008, and the Information technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011.[100]

It lists the grounds for processing of personal data on various grounds, including, but not limited to consent. It follows the GDPR in prescribing limited collection, purpose and storage as well as the notice-based consent model. An interesting aspect is the recommendation to include notices in multiple languages, if possible. It creates the need for data audits, and under Article 35, allows the auditors to assign a rating to data fiduciaries, which must be displayed in the privacy notices to the users. There is also a distinction made between personal data and "sensitive" personal data, and differential rules are laid for processing each category. These rules add an extra layer to the processing of sensitive data, on the grounds of "explicit consent", understood as different from consent. The Bill also creates a separate category of "significant data fiduciaries", which are to be designated by the National Data Protection Authority, based on certain criteria including the quantity of data they process. Third-party data transfers are not explicitly mentioned, implying they are permissible under the specified grounds of processing.

A la GDPR, the PDPB sets up a National Data Protection Authority and an Appellate Tribunal. The former would be a monitoring, advisory, regulatory, quasi-legislative and quasi-judicial body, while

[100] EPW Engage, "What Enables the State to Disregard the Right to Privacy?," *Economic and Political Weekly,* January 16, 2019, pp. 7–8, www.epw.in/engage/article/what-enables-state-disregard-right.

the latter would be an adjudicating authority vested with the powers of a civil court. The Bill does subject the Authority itself to data collection and processing obligations in cases where it processes personal data. What is worth noting is the blatantly missing details and specifications under the two chapters establishing these bodies, most of which are left to be specified at a later date by either the Parliament or the executive branch. This rightly brings into question the autonomy of the Authority, which is essential to perform its functions.[101] The Bill also mentions the need to implement privacy by design, data protection impact assessments, and data breach notifications. It requires data fiduciaries to appoint Data Protection Officers (DPO) to facilitate compliance. Even the penalties have the same ceiling as the GDPR, set at 4% of global turnover in certain cases (Article 69.2). The list of exemptions follows a similar pattern for national security, law enforcement, journalistic and research purposes, and also makes exceptions to manual data collected by homes and small enterprises.

None of the above really distinguishes the proposed Indian legislation from the GDPR, which puts a similar emphasis on free data flow, places limits in the form of exemptions to the protection of personal data, and in fact stresses more precisely the obligations of private operators than those of public authorities. The proposed Indian bill has taken some leaves from the GDPR, as noted both by the Head of the European Commission's International Data Flow and Protection

[101] Bruno Gencarelli, "Submission on Draft Personal Data Protection Bill of India 2018 by the Directorate-General for Justice & Consumers to the Ministry of Electronics and Information Technology (MeitY)," *European External Action Service - European Commission,* November 19 2018, https://eeas.europa.eu/delegations/india/53963/submission-draft-personal-data-protection-bill-india-2018-directorate-general-justice_en.

Unit.[102] The European assessment mentions, over several pages, differences, ambiguities or a lack of legal protection in the proposed bill. On the contrary, a recent Carnegie India examination[103] stresses the multiple features emulating the GDPR, but only to conclude that the scheme is not implementable in the Indian context because of the huge costs of compliance and impossibility for India's small and medium enterprises (SME) to fulfill these tasks. The Carnegie study builds on negative predictions and assessments made by European think tanks about the GDPR: out of eight sources however, only one was published after its actual implementation. The United Kingdom's Ministry of Justice and the European Centre for International Political Economy (ECIPE) – a customarily pro-free trade and anti-regulation think tank – are the main sources from 2012 to 2014. So far, these somber predictions – GDP loss because of the GDPR, fall in data flows to and from the United States – have not materialized.

One criticism does stick. Since SMEs in general find it harder to comply, and given their huge prevalence in the Indian economy and the reluctance to fund data compliance officers, fulfilling a GDPR-like legislation does sound difficult in India. The proposed bill has exempted small firms from most of the obligations, but with high requirements. They must fulfill all these conditions: process data manually with an annual turnover of less than 200,000 INR,[104] not collect data on more than 100 data principals, not collect it for disclosure to other entities, and most importantly, process data manually. This essentially means India's informal sector.

---

[102] *Ibid.*

[103] Anirudh Burman, "Will a GDPR-Style Data Protection Law Work For India?," *Carnegie India,* May 15, 2019, carnegieindia.org/2019/05/15/will-gdpr-style-data-protection-law-work-for-india-pub-79113.

[104] Equivalent to 2500 euros.

While this specific evaluation underlines the similarities with the GDPR and criticizes some of that approach, the European Commission's comments focus on ambiguities and gaps in the proposed bill. Public authorities can exempt data processors from requirements without any other justification than "any law made by Parliament or any State legislature." The Central Government can issue directives to the Data Protection Authority "in the interest of the sovereignty and integrity of India, the security of the State, friendly relations with foreign states and public order." Data processing for law enforcement and national intelligence purposes is bound only by very general requirements. There is little recourse left to individuals and companies on DPO decisions so long as they have been made "in good faith." The right of access to one's own data is limited to "a brief summary."

More Than Just PDPB: Other Legislations and Proposed Legal Texts

The precedents of the PDPB in the data protection sphere would be Section 43-A and Section 72-A of The Information Technology (Amendment) Act, 2008, which laid down compensation for negligence in the processing of sensitive personal data or information (SPDI) and punishment for disclosure of personal information. SPDI were further specified by the IT (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011. However, the proposed PDPB expands the scope for sensitive personal data from its predecessor, the IT Rules, to include official identifiers, information about an individual's sex life, genetic data, transgender or intersex status, caste or tribe. The Rules also require privacy policies, while the requirement for these in the PDPB remains unclear – it only specifies the need for notifications.

In addition, there are certain sector-specific regulations for data collection and processing. For example, the Reserve Bank of India has the competence of a regulatory authority for financial data and has issued data processing rules for the sector. In 2018, the Ministry of Health and Welfare proposed the Digital Information Security in Healthcare Act (DISHA), which outlines, as the name suggests, rules for data collection and processing in the health sector. This bill will be further discussed in a later section.

Like DISHA, a number of other proposed texts involve data privacy in some manner. The 2018 Draft Information Technology [Intermediaries Guidelines (Amendment) Rules apply to any person who, on behalf of another person, receives, stores or transmits an electronic message or provides any service with respect to that message. They must publish rules, regulations, privacy policy and user agreement to inform users on the restrictions to access or use an intermediary's resource(s). Another text in the pipeline, the National E-Commerce Bill, allows the Indian government to access source code and algorithms, while prohibiting third-party sharing of sensitive data, even with consent.[105]

From the Bottom Up – A Sea of Apps

Generally, control over apps is sketchy. The Indian online market is swamped with Chinese and American apps. TikTok, the Chinese video sharing app that is wildly popular throughout the country, explicitly states it "cannot guarantee the security of your information

[105] India's Data for India's Development, "Draft National E-Commerce Policy", February 23, 2019, https://dipp.gov.in/sites/default/files/DraftNational_e-commerce_Policy_23February2019.pdf.

transmitted through the platform", and has been involved in social sharing of rural mob attacks. One press report states that "20 Chinese video apps dominate the mobile entertainment network of tier-2 and tier-3 cities, mostly thanks to titillating videos, suggestive notifications, risqué humor and raunchy content."[106]

Chinese apps are not alone in benefiting from a lack of enforced regulations on privacy. Google Pay and WhatsApp Pay privacy policies in India, although the latter is yet to be rolled out, state that they share data with third parties, as do PayTM and PhonePe (owned by Flipkart).[107] Twitter and other social media sometimes roll out innovations in India because the regulatory environment is less constraining. With concerns rising about the data security of Chinese apps as well as the American GAFA, internally as well as globally, India seems to be looking towards policy tools to navigate the different data protection regimes in these countries.

Cross-Border Data Flow and Data Sovereignty

One of these tools is data localization. There is a legislative effort in favor of data localization in the name of sovereignty and security that is also often judged to be a front for support to local industry and companies. The PDPB has a whole chapter dealing with

[106] Economic Times Online, "Are RSS's Fears about Tik Tok True? Here's What You Should Know," *The Economic Times,* February 19, 2019, economictimes.indiatimes.com/articleshow/68066972.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.

[107] "WhatsApp Legal Info," *WhatsApp,* February 5, 2018, https://www.whatsapp.com/legal?doc=payments-in-privacy-policy&version=20180205, and Shrutika Verma, Mihir Dalal, "WhatsApp May Be Sharing Your Payments Data with Facebook," *Livemint,* April 10, 2018, https://www.livemint.com/Industry/VmupcMWS2ZbVssXuIInP2J/WhatsApp-may-be-sharing-your-payments-data-with-Facebook.html.

cross-border transfer of data which is very clear on this. All personal data that falls under this bill must have at least one copy stored in India, but for "critical" personal data (a category yet to be defined by the Central Government), the data must be stored only within India. For the transfer of this kind of data across borders, the Bill prescribes an environment similar to that of the GDPR, i.e., adequacy-based transfer tools. In another case of the nativist instinct, the draft e-commerce policy discussed earlier is subtitled "India's data for India's development." It mandates having all data stored in data centers and on server farms in India, giving companies three years to comply.

Internationally, the biggest controversy is indeed over data localization – an issue that is primarily economic, even if it has potential implications for data security and privacy as well. The decision reflects a broad economic concern, and a narrower security angle. Most controversially, Mukesh Ambani, who already benefited from the government's push for a universal mobile network, has railed against "data colonialism" and urged to "migrate the control and ownership of Indian data back to India — in other words, Indian wealth back to every Indian."[108] In April 2018, the Reserve Board of India (RBI) ordered companies to store all financial data in India, in order to ensure full supervisory access, with only six months for implementation; a stance it recently reiterated in June 2019 after the government requested reconsideration.[109] The European

[108] Mahesh Langa, "Mukesh Ambani Urges Modi to Take Steps against Data Colonisation by Global Corporations," *The Hindu,* January 18, 2019, https://www.thehindu.com/news/national/mukesh-ambani-urges-modi-to-take-steps-against-data-colonisation/article26025076.ece.

[109] PTI, "RBI to Examine Concerns over Data Localisation Rule: Government," *The Economic Times,* June 18, 2019, https://economictimes.indiatimes.com/news/economy/policy/rbi-to-review-data-storage-rules-for-payment-firms-government/articleshow/69838249.cms?from=mdr.

Commission criticizes the whole policy as one hindering the free flow of data, failing to enhance data security, and essentially being a case of protectionism. It also notes a paradox inasmuch as India "is already a top world leader in the data processing industry,"[110] and cites the risk of retaliatory measures by others. Business critics more often cite the cost – including a drain on India's expensive electricity supply – of investing in massive data servers under adverse climatic conditions.

Big Government, Big Data

Alongside the data sovereignty issue comes the issue of excessive state powers. In fact, one of the key problems with the PDPB, as the European Commission also highlighted in its comments on the Bill, is the power it gives the Central Government to specify certain key clauses. The government can decide the conditions of employment and funding for the Data Protection Authority (Articles 50, 56, and 57), appellate tribunal (Articles 79-82), and define critical data (Article 40.1). For the authority to be independent, it is pertinent to have autonomy, at least over financial matters. For a 62-page legal text, it is problematic to leave so many issues unspecified. The draft intermediary guidelines reflect the same trend. According to Article 3(5), "When required by lawful order, the intermediary shall, within 72 hours of communication, provide such information or assistance

[110] Bruno Gencarelli, "Submission on Draft Personal Data Protection Bill of India 2018 by the Directorate-General for Justice & Consumers to the Ministry of Electronics and Information Technology (MeitY)," *European External Action Service,* September 29, 2018, https://eeas.europa.eu/delegations/india/53963/submission-draft-personal-data-protection-bill-india-2018-directorate-general-justice_en.

as asked for by any government agency…", pertaining to certain specified conditions that follow.[111]

Privacy activists in India have, in recent years, raised a number of concerns about the government's actions, starting from the previously discussed case of the Aadhaar card. There have recently been efforts to streamline the digital data collected through Aadhaar - which qualifies as sensitive personal data under the proposed regime – first in the form of Aadhaar-based authentication and Aadhaar-based Know Your Customer services, and now, as a pool of open Application Programming Interfaces called India Stack.[112] Cases of government actions affecting internet freedom are aplenty, as well. WhatsApp is under pressure to identify and stop mass messages which often are about fake news or encouraging violence: the company says this would violate its encryption pledge. And Facebook, which has 260 million users in India, has in 2015 removed the largest amount of content (across all its services) worldwide at the request of the Indian government.[113] Paytm, the most popular e-wallet in the country, has also surrendered data to authorities on one public violence event.[114]

[111] Ministry of Electronics and Information Technology, "Comments on the (Draft) Information Technology [Intermediaries Guidelines (Amendment) Rules], 2018," December 24, 2018, p.3, https://meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf.

[112] "FAQs - IndiaStack," *IndiaStack*, 2016, https://indiastack.org/faq/.

[113] Christina Medici Scolaro, "Facebook Blocks More Content Here than in Any Other Country," *CNBC,* November 13, 2015, https://www.cnbc.com/2015/11/13/facebook-blocks-more-content-here-than-any-other-country.html.

[114] Madhulika Srikumar, "This Isn't Just About Paytm – Laws on Government Access to Data Need to Change," *The Wire,* May 28, 2018, https://thewire.in/law/paytm-data-theft-cobrapost-sting.

Veering Towards China's Digital Model

The policy direction taken by India is not guaranteed. Modi himself has also courted the largest American companies with a visit to Silicon Valley and a plea to help India become an internet powerhouse.[115] Under the influence of the 2017 Supreme Court ruling, the future data protection bill has decidedly veered towards a GDPR-style legislation, with the lack of legal recourse for individuals as the main difference. But other legislations go in an entirely different direction, that of China, emphasizing national security over free data flow. This is not entirely the case for data localization – if the requirements are close to those enacted by China, there are not the stringent limitations on cross-border transfer that are in place in the Chinese case. But the right of the state to obtain personal data via operators is almost as open-ended as it is in China, and responsibility is being placed explicitly on intermediaries (telco companies or social media platforms). The proposed Intermediaries Guidelines reflect this.

On the basis of an earlier 2011 rule, a wide domain has been retained by the Draft Intermediaries Guidelines, including for instance content that "threatens the unity, integrity, defense, security or sovereignty of India, friendly relations with foreign states, or public order or causes incitement to the commission of any cognizable offence or prevents investigation of any offence or is insulting any

[115] Vindu Goel, "Narendra Modi, Indian Premier, Courts Silicon Valley to Try to Ease Nation's Poverty," *The New York Times,* September 27, 2015, https://www.nytimes.com/2015/09/28/technology/narendra-modi-prime-minister-of-india-visits-silicon-valley.html?action=click&module=RelatedCoverage&pgtype=Article&region=Footer.

other nation."[116] Public health and safety, and critical infrastructures have been added. The Draft strongly supported by the telecom giant Jio's owner Mukesh Ambani, while large foreign companies like Microsoft and Google complain. Microsoft – whose Hyderabad-born chief executive, Satya Nadella, is a business icon in India – says that filtering the full range of content requested by the government would not only violate privacy and freedom of expression, but would also be so challenging that "the cost of even attempting compliance will be prohibitive."[117] The draft national e-commerce policy was also initially envisaged to include data localization for e-commerce data, a provision removed only after comments from the industry. Its inclusion in the scope of the PDPB now lies with India's Ministry of Electronics and Information Technology (MeitY) to decide.[118]

The Limits of an India-EU Comparison

The objections from the European Commission should be qualified. India's Data Protection Bill is the law of a federation that has full sovereignty. The EU does not have this – and the GDPR consequently does not address national security, public order and also leaves exceptions in many areas of "public interest." It is true that the GDPR

[116] Ministry of Electronics and Information Technology, "Comments on the (Draft) Information Technology [Intermediaries Guidelines (Amendment) Rules], 2018," December 24, 2018, p.2, https://meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf.

[117] Vindu Goel, "India Proposes Chinese-Style Internet Censorship," *The New York Times,* February 14, 2019, https://www.nytimes.com/2019/02/14/technology/india-internet-censorship.html.

[118] Anandita Singh Mankotia, "MeitY May Not Include E-Commerce Data in Privacy Bill," *The Economic Times,* August 29, 2019, https://economictimes.indiatimes.com/news/economy/policy/meity-may-not-include-e-commerce-data-in-privacy-bill/articleshow/70884990.cms?from=mdr.

has provided for elaborate appeal and review mechanisms at the state level, in cross-country cases and at the EU level. But areas in which the EU is not competent are left out. Once this has been noted and the exemptions in the European regulation are factored in, the Indian data protection draft bill appears to be less divergent from the GDPR.

However, the government's rights under law remain expressed both vaguely and widely in the PDPB, making the recourse from an individual to a legal process for redress quite difficult. The government's access to private data remains largely allowed under a 1996 Supreme Court ruling[119] that focused on telephone tapping. This, despite having appropriate legal procedures, is also facilitated by an overburdened oversight system.[120]

The strongest differences are elsewhere. While Europe has a limited policy of support to digital companies, mainly through subsidies, the Indian government is pro-active: Digital India, Startup India, Skill India, and the India Innovation Fund all serve this purpose. India has taken a leaf from China's industrial and technological policies. This shows up in several areas: massive facilitation to create, Jio, whose owner Ambani is a major business supporter of Mr. Modi;[121] a push for data digitalization across wide sectors of government, usually starting from the initial Aadhaar platform. Aimed

[119] Supreme Court of India, People's Union Of Civil Liberties... vs Union Of India (Uoi) And ANR. (December 18, 1996).

[120] Zubin Dash, "Do Our Wiretapping Laws Adequately Protect the Right to Privacy?," *Economic and Political Weekly* 53, no. 6 (November 28, 2018): 7–8, https://www.epw.in/engage/article/can-government-continue-unhindered-wiretapping-without-flouting-right-privacy.

[121] Simon Mundy, "India: The Creation of a Mobile Phone Juggernaut," *Financial Times,* October 2018, https://www.ft.com/content/4297df22-bcfa-11e8-94b2-17176fbf93f5.

at creating a national ID system, Aadhaar, based on an iris scan and fingerprints, crossed the 1 billion registered users mark in April 2016. At this point, Nandan Nilekani, the founder of Infosys and first chairman of UIDAI (Unique Identification Authority of India) hyped a "600 billion USD market capitalization opportunity" to extend its use to payment systems.[122] The effort is particularly aggressive in using biometric techniques. New payment systems include thumb recognition features – "your thumb is your bank", explained Mr. Modi.

It must be noted that most of the legal texts discussed in the Indian case are yet to be approved by the Parliament, upon which the outcome of the Indian data protection regime is conditional. India's legislation is thus caught between international and domestic web market actors, a Constitution and its judges who have proven to be protective of privacy rights and a government that sees issues in terms of modernization and efficiency of governance. It is a battleground for both privacy issues and sovereign control of data versus free flows. It remains to be seen which way India chooses to go. India's undecided status is well described in a comparative study that puts the country close to surveillance states such as China or Russia, but yet notes that the coming legislation might reverse much of the situation – if it is indeed implemented.[123]

[122] Nandan Nilekani, "India Financial Sectors," *Credit Suisse,* June 29, 2016, https://research-doc.credit-suisse.com/docView?language=ENG&format=PDF&document_id=1062747711&source_id=emcsplus&serialid=Wm0zJuKszkmbCwRYV7h

[123] Paul Bischoff, "Surveillance States: Which Countries Best Protect Privacy of Their Citizens? - Comparitech," *Comparitech.Com,* October 15, 2019, https://www.comparitech.com/blog/vpn-privacy/surveillance-states/.

V

CHINA, THE SURVEILLANCE STATE WITH SOME PRIVACY CONCERNS

With China, we enter a different world. It is one where there are points in common regarding government control with some dispositions currently envisaged by the Modi government in India, but absolutely no point of contact with GDPR style regulation, except for some figures of speech. China's paramount leader, Xi Jinping, regularly emphasizes the importance of putting China at the forefront of digital and artificial intelligence developments, with almost messianic overtones: It will "fulfill the steady increase in the people's good life."[124] Under his watch, several factors merge to produce the largest and most integrated scaling effects on the planet. However, the most important feature predates Xi's mandate: the Chinese internet functions in practice as an intranet. There is no foreign telco network. China-to-China data never leaves the country. Outside traffic only passes through a few checkpoints that can be shut down.[125] This, of course, is totally different from the heavily integrated Indian web, but China's firewall model is being emulated by Russia.[126]

[124] Xi Jinping, "Message from Xi Jinping to the First China Digital Construction Summit 满足人民日益增长的美好生活," (March 22, 2018).

[125] Catalin Cimpanu, "Oracle: China's Internet Is Designed More like an Intranet," *ZDNet,* July 30, 2019, https://www.zdnet.com/article/oracle-chinas-internet-is-designed-more-like-an-intranet/.

[126] Andrew Roth, "Russia's Great Firewall: Is It Meant to Keep Information in – or Out?," *The Guardian,* April 28, 2019, https://www.theguardian.com/technology/2019/apr/28/russia-great-firewall-sovereign-internet-bill-keeping-information-in-or-out.

The Long-Term State Push for All Digital Industries

Within this closed sphere, long-term programs and massive subsidies have created an infrastructure of 4G mobile phones – at last count 1,56 billion mobile phone subscriptions existed, or more than one per person.[127] This also dominates internet traffic, creating opportunities for use at any moment, anywhere. China's backward cash and state banking system has suddenly been superseded by the world's largest mobile payment system. In 2018, mobile payment platforms registered an astounding 60 billion transactions for a claimed amount of 41 trillion dollars USD, and with an explosive growth rate of 37%.[128] Electronic payments in China already reach 25% of commerce, as opposed to 11% in the United States currently – European usage differs widely across member states. The universal use of scanned QR codes – down to beggars and using toilet paper in public places – also made this possible. It also ensures that data is recorded in a single transferable format. Such is the concentration of China's digital business that two telcos, China Mobile and China Unicom, dominate the entire mobile phone business, while two platforms, Alipay and WeChat Pay, dominate 90% of the mobile payment industry. This concentration is also notable across sectors. Alibaba, among the top ten global companies, is no longer the e-commerce platform that made its initial reputation. It is an ecosystem of platforms ranging from eight wholesale and detail commerce (both domestic and global), five media and entertainment companies, two financial companies (520 million customers) that include Alipay and a facility for lending to small and

[127] Yu Xiaoming, "Govt to Further Boost Advanced Manufacturing, Innovation and Competitiveness," *Chinadaily.com.cn,* March 5, 2019, http://www.chinadaily.com.cn/a/201903/05/WS5c7e0b0fa3106c65c34ecdb1.html.

[128] People's Bank of China, "Overall Situation of the Payment System in 2018 2018年支付体系运行总体情况," March 20, 2019, p.4, http://www.gov.cn/xinwen/2019-03/20/content_5375401.htm.

medium firms, navigation, delivery and a "life search" engine providing local services, a logistics platform,[129] and a starting health insurance business that is slated to take advantage of a very low penetration rate for this type of product in China. As such, Alibaba gathers huge amounts of real-time data from users and businesses alike across most of their daily activities. Its arch-rival, Tencent, possesses similar penetration through the horizontal expansion of its WeChat messaging platform, which also has 100 to 200 million international users. Overall, China's state support for Internet+ policies has greatly favored the so-called BAT – Baidu, Alibaba and Tencent – granting them near data monopoly in their respective market with regard to web searches, online transactions, and social media. This goes hand-in-hand with a very close cooperation afforded to public administrations: much public data is privatized, but the government has, as we shall see, unlimited access. Interestingly enough, the three BAT companies are actually holding companies based in the Cayman Islands, relying on contracts with their China-based subsidiaries to draw dividends.

The Issue of Technological Interdependence

One should not consider that possessing big data implies using it efficiently, whether in terms of analytics or tailored products. The topic of China's real advance in analytics is hotly disputed. While some experts such as Kaifu Lee,[130] who also have skin in the game, extoll China's prowess, there are occasional signs that even the biggest platforms can be dependent on software made in America

[129] Ming Zeng, "Everything Alibaba Does Differently — and Better," *Harvard Business Review,* August 21, 2018, https://hbr.org/2018/09/alibaba-and-the-future-of-business.

[130] Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt, 2018)

or elsewhere. Alibaba, for example, has a partnership with Salesforce, America's leading provider of customer relation management (CRM), to provide cloud-based technologies to its own clients. It signed a similar partnership with AXA, the European insurer, to provide tailored insurance products to its e-commerce customers, Chinese SMEs as well as travelers using Alipay outside of China.[131]

Yet, the example of Alibaba, a company that does not hesitate to strike deals with potential competitors at home and abroad, cuts both ways. Alibaba presents itself as the "exclusive provider" of Salesforce for Greater China, even though it recognizes the power of Salesforce's CRM solutions.[132] The company has developed an app that can "process auto insurance claims in seconds whilst assessing exterior vehicular damage and displaying vehicle damage information to users, including where to repair the vehicle."[133] It is developing health apps that combine diagnosis tools with a search for price bids to fill their prescriptions, payment systems for state hospital patients, which also allow the company to offer second opinions or prescription bids, and ultimately an integrated diagnosis and payment tool. It is providing blockchain technology to manage patient records and prescriptions across provincial borders.[134]

[131] AXA, "AXA, Alibaba and Ant Financial Services Announce Global Strategic Partnership | AXA," *AXA.com,* July 29, 2016, https://www.axa.com/en/newsroom/press-releases/axa-alibaba-ant-financial-services-announce-global-strategic-partnership.

[132] Tom Brennan, "Alibaba Now Exclusive Provider of Salesforce CRM in Greater China," *Alibaba Cloud Community,* July 25, 2019, https://www.alibabacloud.com/blog/alibaba-now-exclusive-provider-of-salesforce-crm-in-greater-china_595141.

[133] The Digital Insurer, "Alibaba - The Digital Insurer," *The Digital Insurer,* November 10, 2018, https://www.the-digital-insurer.com/cif-alibaba/.

[134] Michael O'Dwyer, "Alibaba - The Digital Insurer," *The Digital Insurer,* Undated, https://www.the-digital-insurer.com/cif-alibaba/.

Data Oligopolies

A major difference stems from this portrait and could be applicable to China's other digital giants: their activities, and therefore their data resources, cut across many sectors. The issue of third-party transfer of data, which is so important in the design of America's, Europe's or India's privacy rules, matters less in the Chinese case, where a few platforms form a data oligopoly. Nonetheless, they do not neglect income from re-selling to third parties, as we shall see. WeChat, with 1,1 billion users in China, 900 million for its payment system and 100 million for its financial products, is another case in point. No non-Chinese platform can claim such ubiquity and breadth of services. It is not surprising that Mark Zuckerberg, Facebook's founder and principal owner, talks of moving away from an advertising-based model (with many third-party users or consumers of the company's data), to other services "including calls, video chats, groups, stories, businesses, payments, commerce, and ultimately a platform for many other kinds of private services."[135] Similarly, Google is expanding payment services and investing into AI and cloud services in the health sector.

The point here is that even before one gets to the issue of China's massive surveillance state and its digital arms, its technically private platforms aggregate more personal data than in any other society. In the construction of a national integrated data bank for social credit applied to companies, we find Alibaba, Tencent and Huawei to be key actors. Nor is it confined to these platforms, as China has also seen a host of start-ups grow into new sectors – facial recognition

[135] Li Yuan, "Mark Zuckerberg Wants Facebook to Emulate WeChat. Can It?," *The New York Times,* March 7, 2019, https://www.nytimes.com/2019/03/07/technology/facebook-zuckerberg-wechat.html#click=https://t.co/q1vhufRaCi.

being the best known. The issue of privacy should be paramount. The top business leaders in the digital sector take different attitudes to this. In the words of Jack Ma, founder of Alibaba: "Europe does not have a big internet company, because it has way too much legal system (…) Internet is still at the initial stage, and we are already talking about the issue of privacy and security. Trust me, we will be able to solve the problem, and if not, our kids will."[136] Pony Ma, founder of Tencent, is more cautious: "Data cannot be aggregated without rules. Communication, social exchange and consumer behavior data must not be aggregated, or this will bring catastrophic consequences".[137] He has called for unified rules protecting internet users' data privacy during China's two parliamentary sessions of March 2019.[138]

Regulation Is About Data Security First

But the general discussion about privacy is more focused on data security and the overall need for regulation than on ensuring privacy. In any case, oversight is split between multiple and different authorities – the Ministry for Public Security, the State Administration for Market Regulation (SAMR) and the Cyberspace Administration

[136] Sina.cn, "Ma Yu: The Combination of the Internet of Things and Big Data is the Future 马云：物联网和大数据的结合才是未来," *Sina.cn,* September 10, 2017, https://tech.sina.cn/it/2017-09-10/detail-ifykuffc4789377.d.html?cre=tianyi&mod=wtech&loc=1&r=25&doct=0&rfunc=0&tj=none&tr=25&vt=4&pos=18.

[137] Zhang Chao, "Ma Huateng Answers : Tencent Does Not Dream. Social Exchange, Communication "Data Should Not Be Aggregated without Rules马化腾回应‘腾讯没有梦想：社交、通信‘数据不能任意打通,'" *Shidai Caijing,* November 9, 2018, https://tfcaijing.com/article/page/8a9eaf0566e21b6e0166f3e81bb11c44.

[138] Xinhua, "Ma Huateng: Collection of Private Information Through Big Data Should Not Be Too Complete 马化腾：大数据收集隐私信息不宜太全," *Xinhuanet.com*, March 4, 2017, http://www.xinhuanet.com//politics/2017-03/04/c_1120566998.htm.

of China (CAC), and at least nine other government level entities are involved in national regulation. A few voices call for a GDPR style comprehensive approach to "establish a one stop, unified law enforcement department to be responsible for personal information protection"[139] while stipulating punishments for violations. Experts and official sources alike have a field day, of course, in highlighting the privacy breaches that happen elsewhere. In an article on China's status as a big data power, the People's Daily notes that personal data leakage is on the rise internationally. More ominously, it notes that 96,6% of Android apps installed in China seek access to personal data, and that 25,3% of them have cross-border access to this personal data.[140]

Third Parties Piggyback the Data Oligopolies

A detailed study – done in cooperation with Microsoft and international scholars – delves deeper into the issue. On China's main platforms providing access to third-party apps – Baidu, Tencent and Wandoujia – there is wide-spread launch of other apps in the background, without users being aware of even having previously used these apps. On average, each app launches 76 other apps. On a total of 800 apps running on 1520 devices (mostly smartphones), 27,1% of the energy is consumed by these undetected launches. The apps form clusters as visible in the figure below for Wandoujia:

[139] Cui Xiankang, Han Wei, and Ren Qiuyu, "Proposed Guidelines Highlight China's Fragmented Protection of Online Privacy - Caixin Global," *Caixinglobal.com,* May 9, 2019, https://www.caixinglobal.com/2019-05-09/proposed-guidelines-highlight-chinas-fragmented-protection-of-online-privacy-101413683.html.

[140] Liu Miao, "China's 'Big Data' Is Not Only About "Data Size 中国大数据，不只‘数据大'," *People's Daily Overseas Edition,* July 9, 2018, http://www.gov.cn/shuju/2018-07/09/content_5304898.htm.

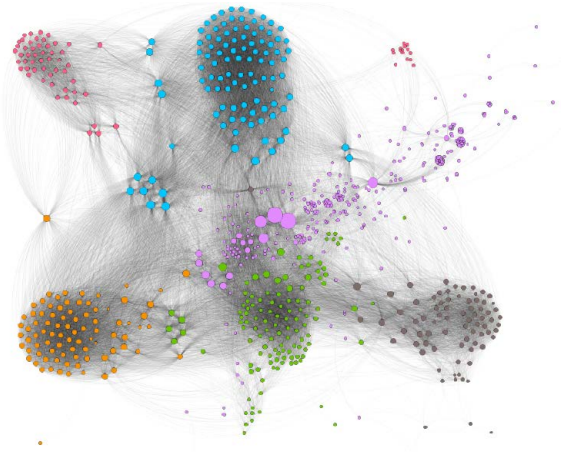

Figure Source: Mengwei Xu et al., "AppHolmes: Detecting and Characterizing App Collusion among Third-Party Android Markets," Microsoft Research, April 3, 2017, p. 148, https://www.microsoft.com/en-us/research/publication/appholmes-detecting-characterizing-app-collusion-among-third-party-android-markets/.

These background apps often require sensitive permissions endangering personal data. Ironically, part of the problem lies with the forced unavailability of the Google search engine in China: apps use push service on Android (another Google product…) to compensate for this gap.[141]

According to a survey conducted by the China Consumers Association in 2018, 85,2% of interviewees experienced some kind of data leak,

[141] Mengwei Xu et al., "AppHolmes: Detecting and Characterizing App Collusion among Third-Party Android Markets," *Microsoft Research,* April 3, 2017, https://www.microsoft.com/en-us/research/publication/appholmes-detecting-characterizing-app-collusion-among-third-party-android-markets/.

but one third of them decide to "swallow it and accept the bad luck."[142]

Start-Ups, The Surveillance Front Line

Finally, the BAT companies' oligopoly on data does not prevent China to have a lively and well-supported scene for start-ups. Some of it, it may be argued, resembles a bubble: 27 start-ups focus on AI. But new contestants do rise. Bytedance is now the highest valued start-up in the world, owning TikTok (the international version of Douyin in China), a video lip-sync app that is wildly popular all over the world with children and teenagers, as previously seen in India. It is now taking traffic away from Alibaba and Tencent. In a good example of what is becoming the global splinternet, its privacy policies differ according to market. By contrast, at least some American platforms have announced they would implement GDPR rules across the board, and not only in Europe. European users under the GDPR, as well as Indian users, can access their personal data from TikTok if they care to. Such was not the case in the United States until February 2019, and the data could be transferred to servers in China. There is actually nothing unusual regarding international data transfer – except that this is highly personal data caught at the most sensitive age, with personal identification, and an everlasting memory in China.[143] Many other companies in China are developing AI apps,

[142] Cqn.com.cn, "China Consumers Association released "Investigation Report On the Personal Information Disclosure of Apps" 中消协发布《App个人信息泄露情况调查报告》," *Cqn.com.cn,* August 29, 2018, http://www.cqn.com.cn/pp/content/2018-08/29/content_6213791.htm.

[143] David Carroll, "TikTok Might Be a Chinese Cambridge Analytica-Scale Privacy Threat," *Quartz,* May 7, 2019, https://qz.com/1613020/tiktok-might-be-a-chinese-cambridge-analytica-scale-privacy-threat/.

running for example on facial recognition that have multiple surveillance uses, and not only for the state. Hanwang, founded in 2014, provides facial and biometric recognition services and optical character recognition, as well as air quality monitors and purifiers. One of their apps, running with Hikvision cameras, offers to schools a "Class Care system" that watches individually every student's attitude before giving it a weekly score. This is one of the everyday consequences of the government's Next Generation Artificial Intelligence Development Plan (NGAIDP). The plan aims to incorporate AI in virtually all aspects of life, including medicine, law, transportation, environmental protection, and what it calls "intelligent education."[144] The ubiquity of surveillance cameras – estimated at 176 million in 2017, and projected at 626 million by 2020,[145] is a big advantage for the collection of data. The uses are now everywhere across China.[146] This is only the everyday societal consequence of a vision that includes the sinister implementation of AI and facial recognition programs against Xinjiang's entire population.[147] In an example inadvertently exposed, one such system

[144] Xue Yujie, "Camera Above the Classroom," *Sixth Tone,* March 26, 2019, https://www.sixthtone.com/news/1003759/camera-above-the-classroom.

[145] According to widely cited statistics gathered by Statista in 2017. Source: Statista Research Department, "China: Surveillance Camera Installation 2017-2020," *Statista*, 2019, https://www.statista.com/statistics/879198/china-number-of-installed-surveillance-cameras/.

[146] For various examples, see: Julie Zaugg, "En Chine, La Vie Sous l'Oeil Inquisiteur Des Caméras," *Les Echos,* March 7, 2019, https://www.lesechos.fr/tech-medias/hightech/en-chine-la-vie-sous-loeil-des-cameras-997774.

[147] HRW, "China's Algorithms of Repression Reverse Engineering a Xinjiang Police Mass Surveillance App," *Human Rights Watch*, May 1, 2019, https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance.

was gathering daily detailed individual information on 2,56 million people in Xinjiang, based on surveillance cameras.[148] Fascinatingly, Xinjiang is designated by the Ministry of Information Technology as the center of « pilot projects » for big data integration and Artificial Intelligence in 2020.[149]

China's Social Credit – A Hydra Larger Than Life

Since 2014, no innovation in control has sparked as much comment as China's "social credit system".[150] Part of the reason is that it was officially hyped in China as a path to trust by enforcing a reward/punishment system. One seldom sees in China a mention of the evident imitation of scoring techniques that are in wide use in market economies – a credit score in the United States, for example, is mandatory not only for credit cards, loans and insurance, but also for renting property. Nor are inroads into privacy unique to China. Massive and publicly available data bases in the United States provide information on any individual, including traffic fines, and allow for example one to locate known sex offenders in your neighborhood.

[148] Catalin Cimpanu, "Chinese Company Leaves Muslim-Tracking Facial Recognition Database Exposed Online," *Zdnet.com*, February 17, 2019, https://www.zdnet.com/google-amp/article/chinese-company-leaves-muslim-tracking-facial-recognition-database-exposed-online/?__twitter_impression=true.

[149] Ministry of Industry and Information Technology of China, "Notice of the Ministry of Industry and Information Technology of China, on Applying to the Big Data Pilot Project in 2020 工业和信息化部办公厅关于组织开展2020年大数据产业发展试点示范项目申报工作的通知," November 6, 2019, http://www.miit.gov.cn/n1146295/n1652858/n1652930/n3757022/c7517097/content.html , cited by *EastisRed (Passe Muraille)* n° 38, November 18, 2019, https://eastisred.fr/passe-muraille-n38-semaine-du-11-novembre/.

[150] 社会信用体系 (shehui xinyong tixi).

Furthermore, China goes back a long way in terms of collective surveillance – the "100 household" *(baojia)* system in Imperial times was a system of mutual surveillance. In a Leninist-Maoist context, the entire population was classified according to some 40 categories, depending on a mixture of class origin, personal status and behavior, from good to bad, red to black. The Communist Party of China system has always had *dang'an* – files from all sorts of surrendered or collected information – on every one of its 90 million members (as of 2018). Conversely, at the grassroots, where the state had always lacked presence, the issue of trust has always been paramount. China was by necessity a society based on relationships (guanxi) because this was a way to overcome distrust. The practice of multiple names and voluntary disappearances was also very frequent, absent a reliable nation-wide ID system.

Against this background, social credit scoring is recreating trust without the need for special relationships – and that is generally welcomed by the population, which values public order and discipline against a traditional background of individualism and unreliability.

The decision in 2014 to construct a social credit system by 2020 combines multiple local experiments with the overall holistic goal of "a market improvement of the economic and social order."[151] The peculiarity of China's political order resides in this single sentence: the market serves public order, and it is therefore difficult to separate public and private initiatives, for example consumer credit scoring and a more general categorization and reward or punishment of good/bad behavior in much wider areas.

[151] State Council of the People's Republic of China, "Significant Improvement in Economic and Social Order "Notice of the State Council on Printing and Distributing the Outline of the Construction of the Social Credit System (2014-2020) 经济社会秩序显著好转 "国务院关于印发社会信用体系建设规划纲要（2014—2020年）的通知"", *www.gov.cn,* June 14, 2014, http://www.gov.cn/zhengce/content/2014-06/27/content_8913.htm

Yet, at this point, it is not sure that all or even most of the social credit systems will be integrated in a single nation-wide scheme – multiple voices are raised against this, which in any case would require a huge technological feat to classify, store and protect the data, perhaps in excess of any actual need. Still, the main systems that exist are both impressive and ominous. Some local or pilot initiatives are scary and represent even worse challenges to privacy.

One scheme, created by China's central bank in 2015 with eight entities, including Sesame Credit (developed by Ant Financial, an affiliate of Alibaba) and Tencent Credit Information Co., aggregates data from online behavior and purchases, public agencies and third-party merchants to create instant scoring of individuals. The ratings include criteria of stability, website behavior and the company you keep – your social relationships for instance. By 2015, Sesame Credit alone rated 300 million individuals and 37 million small firms.[152] In turn, credit scoring data is used by other scoring firms – for example, by the BaiHe dating app to assess your partner. A "game" allows friends to compete on their credit scoring. Many local social credit systems also use data from these large nation-wide companies.

But the build-up does not stop there. A national and publicly-run website, Credit China, aggregates scoring data from 44 central departments, 22 provincial platforms and 122 social institutions. All sorts of data – from taxation, to food and drug or environmental protection are shared, along with the "black" or "red" lists of individuals. To fight bureaucratic "slumber", e.g. non-use of the data, third-party companies can access it for a fee. A company,

[152] Ant Financial, "Ant Financial Unveils China's First Credit-Scoring System Using Online Data," *Ant Financial,* January 28, 2015, https://www.alibabagroup.com/en/news/press_pdf/p150128.pdf.

BaiHang Credit, was created with 35% state ownership – the National Internet Finance Association of China – and 8% by each of the original and technically private pilot companies. Thus, private scoring and public-private sharing of the data are achieved, again revealing China's uniqueness.

The best-known public consequence of this integrated network is the ability of China's courts, from data available with the National Public Credit Centre, to deny some services: at the end of 2018, 17,5 million flights, 5,5 million train trips were denied. There were also, according to the People's Daily, black lists for golf courses, high-end hotels and flats, private schools, and financial products.[153] In one extreme case, a Shandong province court ordered phone companies in 2017 to insert automatic warning messages on calls received by untrusted persons.[154] At the other end, good grades for social behavior, including acts of social benefit and general reliability, win reduced prices on many services, all the way to priority lines at hospitals. Scores can determine people's access to public services, welfare, school admission, employment, job promotions and business undertakings. The schemes are said to involve less than 10% of the population in either red or black categories – but that was always a recipe of traditional Maoism that pitted large majorities against targeted minorities and rewarded a minority of "activists." The same system of reward and punishment conducts "to create a more regulated, fair, transparent and

[153] Xue Yuan, "The Release of 2018 Annual Report on the Credit Blacklist 2018 年失信黑名单年度分析报告发布," *www.gov.cn*, February 19, 2019, http://www.gov.cn/fuwu/2019-02/19/content_5366674.htm.

[154] BBC News in Chinese, "From Portfolio to Credit Score, Is China on the Way to an "Orwellian" Monitoring Society 从档案袋到信用评分 中国是否正走向 '奥威尔式' 监控社会," *BBC News in Chinese,* October 17, 2018, https://www.bbc.com/zhongwen/simp/chinese-news-45886126.

predictable legalized business environment."[155] For instance, payment delay will put you on the black list, and both the company and legal representative will face "obstacles".

The variety of experiments and systems in place has led to debates on their limits: should social credit be based on explicit and clear legal criteria, or should it extend to moral judgments that are loosely defined? The right to be forgotten, and even more concretely, how to restore one's credit – particularly in case of mistakes – are objects of discussion. The issue of "rumor spreading", often another name for criticism of authorities, is also prevalent. Some experts call for a nation-wide regulation on social credit – which as of the spring 2019 was only a "class three priority" for China's national legislature, meaning not forthcoming in the near future.[156]

A Diminutive Public Debate on Privacy

Given the wider pattern of "social credit" systems put in place across China, the Xinjiang case can be considered as an experiment with potential implementation elsewhere. But there is little public debate in China on these aspects, nor are there many experts writing on broader aspects of privacy rights. What exists is focused on cases elsewhere, usually in America or from American companies. We have already discussed the case of Android. Huawei's executives

[155] Luo Pan, "Ministry of Commerce: The Construction of the Corporate Social Credit System Will Not Adopt the So-Called Suppression Measures商务部：企业社会信用体系建设不会采取所谓打压措施," *chinanews.com,* August 29, 2019, https://www.chinanews.com/gn/2019/08-29/8941547.shtml.

[156] Zhang Yuzhe and Han Wei, "In Depth: China's Burgeoning Social Credit System Stirs Controversy - Caixin Global," *Caixinglobal.com,* April 1, 2019, https://www.caixinglobal.com/2019-04-01/in-depth-chinas-burgeoning-social-credit-system-stirs-controversy-101399430.html.

abroad occasionally extol GDPR regulations, a part of the company's style of messaging that is completely absent in China.[157] A discussion on privacy in a consumer magazine hinges around Facebook and a Georgetown University expert,[158] another in a popular science journal also revolves around Facebook and Deepmind.[159] It seems that publications in the People's Republic of China cannot find examples drawn from its own digital industries and practices.

An exception however exists concerning data collection for commercial purposes. In 2017, a government-backed consumer NGO in Jiangsu province launched an enquiry against 27 snooping apps scraping consumers' personal data. It then initiated a lawsuit against Baidu, the sole company among the 27 that did not withdraw the feature. The case went to court; Baidu retreated and updated its app. The NGO then withdrew its lawsuit.[160] As we shall see, there can be tensions between regulatory agencies and commercial companies.

Yet, there are many more relevant discussions of digital privacy when it is framed in the context of data security, whether this is about fraud and abuse of data or about national security.

[157] Joy Tan, "Transparency and Privacy Go Hand in Hand," *Linkedin.com,* November 25, 2018, https://www.linkedin.com/pulse/transparency-privacy-go-hand-joy-tan/.

[158] "Facial Recognition, Privacy Protection and the Scientific Challenges, 刷脸，隐私保护与科技的博弈," *Consumer Daily* 126, July, 2015.

[159] Fang Lingsheng, "Identification Systems: The End of Personal Privacy, 识别系统：个人隐私终结者," *World Science*, March 2015, p. 30-37.

[160] Zhang Jie, "Consumer Rights Group Withdraws Complaint against Baidu," *Chinadaily.com.cn,* March 15, 2018, http://www.chinadaily.com.cn/a/201803/15/WS5aaa1535a3106e7dcc141dda.html.

Digital China's Regulatory Framework

The burgeoning cybersecurity and data protection regulation reflects a familiar triangle of tensions between different goals: efficiency, which in this case means faster development of AI and big data applications; the protection of personal data, which is largely perceived as preventing its misuse by private actors; and the overriding concern for national security, which is perhaps the best-known aspect internationally. This is evidenced by the balanced maze of regulation being constructed. Two major difficulties restrict interpretation: it is difficult to hierarchize laws, regulations and standards since much of the "informal" guidance can actually be quite binding. And as always with Chinese legal texts, ambiguities abound and implementation is extremely variable.

Still, the corpus under development is impressive. After a set of administrative measures in 2000 regulating the internet, China's 2017 cybersecurity law has become the overarching law of reference. The law tilts towards "guaranteeing cybersecurity, safeguarding cyberspace sovereignty, national security and public interest", although it also concerns the legal rights of individuals and organizations. It compels network operators to provide support to all security organs safeguarding national security and investigating criminal activities "in accordance with the law." It also asks all operators to "voluntarily contribute" to the security of "critical infrastructures" that are broadly defined: "public communication and information services, power, traffic, water resources, finance, public service, e-government, and other." The inclusion of the term "other" as an escape clause allowing for any extension is a very frequent occurrence in Chinese public law – from the Penal Code to these cyber-related rules. It essentially leaves authorities free to use their

own definition as the occasion may require. “Critical” data, again very broadly and loosely defined, must be stored in China, a provision that drew criticism from foreign firms and operators in China. Data operators must pass periodic security reviews. The law also requires all users to provide real name information. The government can “take temporary measures regarding network communications in a specially designated region, such as limiting such communications”, a black-out that has been applied on occasions to Xinjiang.

The Cybersecurity Law and Personal Data Protection

In principle, the law provides protection to users: network operators “shall strictly maintain the confidentiality of user information they collect”, and users have a conditional right to erasure or correction of their personal data if it has not been gathered legally or if it is erroneous. It defines “personal information”[161] as “all kinds of information, recorded electronically or through other means, that taken alone or together with other information, is sufficient to identify a natural person’s identity, including but not limited to natural persons’ full names, birth dates, national identification numbers, personal biometric information, addresses, telephone numbers, and so forth.”

As it is, China’s cybersecurity law has some common traits with India’s 2018 “Draft Intermediary Guidelines”: the introduction of “critical” data, the ambiguous mention of provisions “according to the law.” Besides opening exceptions in “other” situations, it restricts all critical data to China. India’s restrictions are narrower in scope.

[161] 个人信息 (geren xinyi).

However, he possibility to completely suspend communications is no longer a unique provision – mobile internet data service in Kashmir and other parts of India have been suspended on occasion under Section 144 of India's Penal Code regarding unlawful assembly.

In our triangle, China's cyber law of 2017 tilts very much towards the state's rights, the private operators' obligations, and a few conditional or generic rights for the individual.

The On-Going Regulatory Maze

It was accompanied or followed by a spate of laws, regulations and standards: one can count six systems in different areas (data protection being one of the six) and more than ten new "standards" so far: they join the 240 existing standards set since 2010 on "Information security technology."[162] In our area of reference, they cover cross-border transfers, national intelligence and counter-espionage. Many other regulations exist on given sectors, including finance, banking, e-commerce and consumer protection. Vagueness and ambiguities abound – including in some cases on whether the rules are mandatory or just guidance. In part, this is also due to China's peculiar legal system, where rules are written on the go, and then eventually rewritten or amended.

Most of these rules point in the direction of even more state control. Cross-border rules prohibit external copy of many fields of data stretching to the customary "other circumstances that possibly affect

[162] They are run by the China National Information Security Standards Technical Committee (CNISSTC), and published exclusively in Chinese on its website at https://www.tc260.org.cn/

national security and societal and public interests." The most recent draft law for cross-border transfers no longer requires official assessment for companies with more than 1000 Giga or 500.000 individuals, but it puts greater responsibility on all companies, which must store their digital data for five years. The risk assessment also remains very broad and open-ended. Critical infrastructures now include media, e-commerce, e-payment, search engines, emails, blogs, cloud computing, enterprise systems and big data. Article 7 of China's 2018 National Intelligence Law – widely cited in the debate over Huawei – states that every "organization or citizen shall support, assist in and cooperate in national intelligence work in accordance with the law and keep confidential the national intelligence work that it or he knows." China's Counter-Espionage Laws unsurprisingly carry similar obligations but the 2017 version goes two steps further. It applies to persons in China *and abroad,* while tying with espionage, actions such as "fabricating or distorting facts, publishing or disseminating words or information that endanger state security, or making, distributing or publishing audio-visual products or other publications endangering state security."

The 2018 Personal Information Security Specification – A Watershed

To individuals or to businesses alike, the most important new rule is the 2018 Information security technology — Personal information security specification (PIS).[163] There are two important provisions

[163] National Information Security Standardization Technical Committee, "Information Security Technology — Personal Information Security Specification 信息安全技术个人信息安全规范," *National Information Security Standardization Technical Committee,* May 1, 2018, https://www.tc260.org.cn/upload/2018-01-24/1516799764389090333.pdf

though: as a "specification" or "standard" *(guifan),* it is not law, yet it is often regarded as such by authorities. And it already has been amended and constantly supplemented by new and specific regulations.[164] In particular, a new draft text on Data Security Management seems to go further in the obligations placed on companies to protect personal data.[165] Keeping this in mind, it is important to note that the PIS is both heavily influenced by the GDPR and yet differs in some key aspects. The drafters of the PIS have extracted the "essence" of available documents (OECD Privacy Guidelines, APEC Privacy Guidelines, GDPR, related ISO law, American laws, etc.) and customized them for the case of China.[166] But they also made it clear that, since the beginning, they were aiming for a regulation stricter than in the United States, but not as strict as in Europe.[167]

Similarities have been described by one external observer,[168] while differences are spelled out by one of the experts who contributed to

[164] Yan Luo, "China Releases Draft Amendments to the Personal Information Protection Standard," *Inside Privacy,* February 11, 2019, https://www.insideprivacy.com/international/china/china-releases-draft-amendments-to-the-personal-information-protection-standard/.

[165] The draft measures have been published in CAC's website.

Source: CAC, "Notice of the National Internet Information Office on Public Consultation on the "Data Security Management Measures (Draft for Comment)" 国家互联网信息办公室关于《数据安全管理办法（征求意见稿）》公开征求意见的通知," *Cac.gov.cn,* May 28, 2019, http://www.cac.gov.cn/2019-05/28/c_1124546022.htm.

[166] Sina Technology, "The story behind the issuing of 'Personal Information Security Specification' compromises of 33 experts made the Standard Possible《个人信息安全规范》出台记：33专家博弈炼就标准," *Sina.com.cn*, May 1, 2018, http://tech.sina.com.cn/i/2018-05-01/doc-ifzvpatr7140886.shtml.

[167] *Ibid.*

[168] Samm Sacks, "China's Emerging Data Privacy System and GDPR," *Csis.org*, March 9, 2018, https://www.csis.org/analysis/chinas-emerging-data-privacy-system-and-gdpr.

the drafting of the text.[169] PIS details obligations for user consent, starting from the minimisation of data collection and secondary use, as well as placing third-party operators under security requirements. A category of "sensitive" personal information is created, including "identity card numbers, biometric information, bank account numbers, communication records and contents, property information, credit information, location data, accommodation information, health and physiological information, transaction data, and the Personal Information (PI) of children 14 years of age or under." De-identification is required. This, the obligation for firms handling large amounts of personal data to have appointed data officers, and the provision of penalties, led some to conclude that PIS is on a GDPR framework.

While "consent" is only one of the six legitimate reasons for lawful processing of data in Article 6.1 of the GDPR, Article 41 of the Chinese cybersecurity law has "consent' as a must, but provides a list of exceptions in the PIS. Articles 5.4 and 8.5 list exemptions relating to national security and defense, public safety, public health, and significant public interests, criminal investigation, prosecution, trial, and judgment enforcement, etc.; safeguarding the major lawful rights and interests such as life and property of PI subjects or other persons, and it is difficult to obtain the consent of the PI subject; when necessary to maintain the safe and stable operation of the provided products or services, such as to detect and handle product or service malfunctions; when necessary for the PI controller, as a news agency, to make legal news reports; when necessary for the PI controller, as an academic research institute, to conduct statistical

[169] Hong Yanqing, "Answers and explanations on five points regarding the Personal Information Security Certification 对《个人信息安全规范》五大重点关切的回应和解释," *WeChat,* February 5, 2018, https://mp.weixin.qq.com/s/rSW-Ayu6zNXw87itYHcPYA.

or academic research in the public interest, which also has de-identified the PI when providing academic research or results externally; and finally, "when other situations specified by laws and regulations." Yet, it might still seem that Chinese regulations leave less leeway than the GDPR, as a list of exceptions cannot compete with the overall flexibility given by what is called "legitimate interests" in the GDPR's Article 6.1.

The Big State as an Arbitrator Between Individuals and Companies

Hong Yanqing takes the time to explain why this is not the case, and how PIS dispenses with user consent in many instances. Justification of "legitimate interests" for the GDPR, requires companies to track each stage of its internal procedure, to be ready in case of demand to supply evidence of "legitimate interests." This actually makes the company more accountable. Hong describes a consent requirement that is much looser in the PIS than in the GDPR.[170] The Standard requires explicit consent in case of sensitive personal information, but only authorized consent, a definition which was not provided in the Standard, in other cases of personal information. Hong further explains that "while using the term 'authorized consent', I mean that you are encouraged to adopt explicit consent, but if it is not feasible in reality, implicit consent can be used."[171] The loose

[170] Sina Technology, "The story behind the issuing of 'Personal Information Security Specification' compromises of 33 experts made the Standard Possible《个人信息安全规范》出台记：33专家博弈炼就标准," *Sina.com.cn,* May 1, 2018, http://tech.sina.com.cn/i/2018-05-01/doc-ifzvpatr7140886.shtml.

[171] 洪延青, "Answers and explanations on five points regarding the Personal Information Security Certification 对《个人信息安全规范》五大重点关切的回应和解释," *WeChat,* February 5, 2018, https://mp.weixin.qq.com/s/rSW-Ayu6zNXw87itYHcPYA.

and evolving Chinese regulation is often put into the context of required innovation and economic efficiency. A strict regulation on personal data, such as the GDPR, is not suitable for China considering China's personal data protection capability and the data industry's current development situation.[172] Indeed, efficiency and the interest of companies were reflected in the PIS as the drafting team of the PIS is composed of 33 individuals, which can be broadly divided into two camps: companies and experts.

The debate about this was salient enough that a 2019 revision of the cybersecurity law (now undergoing a comments phase) has gone into two directions: adding an exemption "when related to the obligations of personal information controllers to perform laws and regulation" of the state; but doing away with the exemption of consent "when necessary to sign and perform a contract according to the PI subject's request" and otherwise strengthening the obligations of companies in the process of ensuring data protection. In other words, in our triangle cited above, the third point – the state's ultimate interests in wide data collection – are, once more, the winner. But the protection of personal data also wins at the expense of companies, which are saddled with its implementation. The Chinese state can protect individuals as consumers against predatory commercial interests, although it will not perform the same task against itself.

One cannot end this descriptive attempt without registering two tentative conclusions. First, there is not much to constrain the Leviathan itself – the surveillance state. Second, an abundance of rules, often amended, supplemented or rewritten, coexists with

[172] Hu Wenhua and Kong Huafeng, "The Impact of EU General Data Protection Regulation on China and Its Response," *Computer Applications and Software* 35, no. 11 (November, 2018).

persisting ambiguities at several levels: the distinction between mandatory rules and suggested guidance is tenuous. The law can therefore be applied sparingly. Or, with the added breadth of "other" categories, it can be stretched arbitrarily. The lack of provisions regarding the means for implementation suggests that the law is mainly used as a deterrent. The obligation for companies handling much personal data to employ data officers has even been pared down – in the 2017 Cybersecurity Law, the threshold for this obligation was 500,000 individuals concerned. It went up to one million in the February 2019 draft revision. Although both the GDPR and India's nascent regulation have their limits, none of them come close to China's conjunction of black holes and regulatory maze.

VI

IN-FOCUS: HEALTH DATA AND PRIVACY

Digital data processing, followed by big data analytics and now AI, all have a huge potential to improve health care. Quick diagnosis and predictive tools, wearable devices, image interpretation and machine learning, telemedicine and genetic or behavioral factor analysis hold much promise for this sector. This revolution is as fundamental as the discovery of vaccines and antibiotics was in the past. The investment and revamping of existing data, needed to deliver on those promises, should not be underestimated: a key actor such as Deepmind, with its health division at the forefront of research (now merged into Google Health), is incurring large deficits. Health-related big data represented 153 exabytes in 2013 and is projected to rise to 2,314 exabytes by 2020.[173] Recent applications include lung tumor detection and prognosis, eye-scan and glaucoma, acute kidney injury, but also statistical correlations from large data bases, such as between drinking and the onset of Alzheimer. Electronic medical records (EMR) save time and therefore money. That's if health professionals have easy tools to enter the required data: especially in decentralized public systems, ergonomics is often forgotten. Strikingly, digital innovations revolutionize both advanced medical research and disease prevention while facilitating care at the grassroots.

The implementation of big data and AI also brings threats: the predictive aspects can have fearsome applications for health and

[173] Research and Markets, "Global Big Data in Healthcare Market: Analysis and Forecast, 2017-2025 (Focus on Components and Services, Applications, Competitive Landscape and Country Analysis)," *Researchandmarkets.Com*, March 2018, https://www.researchandmarkets.com/research/wbhh4n/11_45_bn_big?w=5.

life insurance, and more widely for the confidentiality of an individual's medical condition. Whereas insurance and financial credit have rested on the pooling of risk, the availability of medical data and predictive tools, beyond the already prevalent health questionnaires, could individualize risk profiles to the point where insurance loses its very purpose. Health data is only useful if it can be shared among concerned health professionals, medical researchers but also out of necessity with the public or private insurance entities that underwrite medical treatment costs. The risks of hacking and other security leaks are large, especially if data storage is decentralized. The lack of information available to the public, combined with fears of divulgation of data to banks, insurances, employers and even next of kin, can result in reluctance from patients to turn over their data. On the frontline are general medical information websites which are often the first to sell their visitors' data and to deny easy notice and consent process.[174] In what is a classic case of the battle between the sword and the shield, the numerous limitations of anonymization and pseudonymization techniques have already been mentioned.[175] This also leads to the search for a new solution, in the form of simulated patient data, to be discussed in the next section.

GDPR Places Health Data Under the Public Interest Clause

How do our test cases – Europe, India and China – approach the regulation of health-related data and privacy issues? As we shall see,

[174] Martin Untersinger, "Données Personnelles : Les Mauvaises Pratiques Des Sites de Santé," *Le Monde.Fr,* September 4, 2019, https://www.lemonde.fr/economie/article/2019/09/04/donnees-personnelles-les-mauvaises-pratiques-des-sites-de-sante_5506226_3234.html.

[175] Cf. page 36.

much is still under review. Health is perhaps the key digital sector for which specific rules are needed, and where different objectives must be reconciled – protection of sensitive personal data, medical research and improvements in the provision of health care, and financial requirements in an era of soaring medical costs.

For the EU, the GDPR has a very generic approach, although its Article 9 recognizes health as a special category of personal data. It includes "data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status" (Article 4). Member states can impose further limitations to the processing of "genetic data, biometric data or data concerning health." But health data is a prime example where processing "for reasons of public interest" is authorized without the consent of the data subject (Recital 54), excluding other purposes for third parties such as employers or insurance and banking companies. This exemption to consent for reasons of public or legitimate interest is very wide: it covers health status, including morbidity and disability and their determinants, health care needs and resources, the provision of health care, expenditure and financing, and the causes of mortality. The exemption also covers the right to erasure.

However, public and private insurance are treated differently: processing is authorized "to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system" (Recital 52), even without consent (recital 54). But it "should not result in personal data being processed for other purposes by third parties such as employers or insurance and banking companies." It is interesting to note that cost efficiency is recognized as a need for public health systems, but no special

case seems to be made for private health insurance within the larger issue of private companies. Interestingly, Chinese researchers have looked into the financial impact of the GDPR for hospitals, from its adoption in 2016 to the first months of implementation in 2018. The study emphasizes the costs of achieving compliance – but also notes the growing gap between hospitals that are able to provide digital health services (considered more efficient) with those that do not have the capital or human resources. It concludes that in the longer run, only the former will survive in an open environment. Achieving compliance is thus a way to reach a higher degree of performance.[176]

Rather generic on health data protection, the GDPR leaves a lot of room for member states to decide their own rules, supposedly because they fall under the exempted category of "public interest." However, in legislating these exempted categories, member states are allowed to go beyond the GDPR's scope, not under. The Commission's one year after stakeholder reporting exercise repeatedly notes that different national interpretations or rules still pose problems in the health sector. For insurance, rules for dispensing with explicit consent vary from country to country. Pharma companies note that the interpretation of safeguards needed for the processing of research data can still vary from country to country. In the health field, "the use of the specification clauses in GDPR by member states has created considerable hurdles for companies operating cross-border."[177]

[176] Bocong Yuan and Jiannan Li, "The Policy Effect of the General Data Protection Regulation (GDPR) on the Digital Public Health Sector in the European Union: An Empirical Investigation," *International Journal of Environmental Research and Public Health* 16, no. 6 (March 25, 2019): 1070, https://doi.org/10.3390/ijerph16061070.

[177] Multistakeholder Expert Group, "Contribution from the Multistakeholder Expert Group to the Stock-Taking Exercise of June 2019 on One Year of GDPF Application," June 13, 2019, p. 22, https://ec.europa.eu/commission/sites/beta-political/files/report_from_multistakeholder_expert_group_on_gdpr_application.pdf.

France As a Test Case

It is interesting to consider the case of France, because it combines some contradictory features. On the one hand, since 1945 and the creation of a nation-wide insurance system (that is also a regulator and public buyer of drugs and medical-related equipment), there is quasi-universal tracking and recording of medical acts and health related expenses. A single national filing system (SNIIRAM), for pseudonymized expense claims and reimbursements, was created in 1999 and revamped into an even wider health data national system (SNDS) in 2017: this includes several other data banks from hospitals and relative to causes of deaths. The system is often hailed as unique in French sources, because of the range over time and its inclusiveness. Nowadays, it is in fact far from being unique, as the aggregation of health data resources, their digitalization and homogenous treatment are spreading across countries. In fact, the restrictions to use put in place by French law, the data siloes across various institutions and the constraints to use threaten to place France, and singularly medical and pharma research, in a disadvantageous position. Yet, the existing data banks have also attracted the attention of the French regulator in charge of the GDPR for involuntary privacy breaches: insufficient pseudonymization and weak protection of local terminals have been called out.

The early Loi Informatique et Libertés (Information and Liberties Law) (1978) prohibited the processing of health data (without defining it) with important exceptions : where there is express consent by the patient ; processing necessary for preventive medicine, medical diagnosis, provision of healthcare or treatment, or for the management of healthcare services carried out by a member of a medical profession; statistical processing carried out by the National Institute

of Statistics and Economic Studies (INSEE) ; processing necessary for medical research. The French Data Protection Act (2018) limits the processing of biometric, genetic and health data for public interest purposes. A December 2018 public ordnance further defines limits to the treatment of personal data (including health and other sensitive data such as religion or sexual orientation). But it also creates exceptions wide enough that one could drive a truck through. Article 5 of the ordnance lists as exceptions six cases, one of which includes "processing by public authorities in the pursuit of their missions, if treatment is necessary to fulfil the legitimate interests of the processing entity or of a third party, unless the prevailing fundamental rights and freedoms of the individual concerned require personal data protection, notably if the individual concerned is a child." A People's China lawmaker couldn't have written it more convoluted and ambiguous.

The system suffers from known deficiencies, which are partly about process, but also about purpose limitations. Because the data was collected for reimbursement purposes, its medical content is often limited to short categorizations. As in other countries, actual medical information is both siloed and often kept in non-digital form, or in non-standardized digital form. To make this data inter-operable, beyond the individual exchange of medical data regarding single individuals, is a huge task. One step towards this is the generalization of individual shared medical files (DMP) and the creation of a cloud service hosting this highly sensitive data. For the time being, the DMP is more about information files than about a single data format, which limits its wider use. So far, the national shared medical file system has registered 6 million individuals – but the Paris hospital system alone has 10 million registered patients, a discrepancy that shows how difficult it is to combine information into a unique or

coordinated data bank.[178] It seems predictable that single hosting will, in all likelihood, lead in the future to format standardization. In case of epidemics outbreak, standardization will allow for faster intervention. At present, it is striking that Amazon or Google, by tracking searches or orders of over-the-counter drugs, can chart flu outbreaks more quickly and accurately than any medical or epidemiology service.[179]

Another difficulty relates to the conditions of access to this data. Terms differ between health professionals and commercial, medical or insurance companies, a welcomed restriction. But for all, there is at present a necessity to justify access with a single and clearly defined purpose: this defeats the purpose of factor identification through AI – a process which is more akin to a fishing expedition where one does not know where and what the results will be. This is clearly tied to the cultural reluctance for surrendering vital data. The "privacy paradox" works very well to dissipate this reluctance in the daily life of consumers. In the case of health data, where it is harder to identify the immediate and short-term return for the individual of surrendering one's private data, it is less effective. Better information about the use of data could change this bias. Reassurances and education are clearly called for.

At present, French pharma companies therefore complain with some justification that they must turn over to other databases. The United States has huge health data resources, leading to what could be termed an arms race between emerging local privacy laws and

[178] "Les hôpitaux de Paris ont ouvert près de 10 millions de dossiers patients", *Les Echos,* October 28, 2019.

[179] Ali Alessa and Miad Faezipour, "A Review of Influenza Detection and Prediction through Social Networking Sites," *Theoretical Biology and Medical Modelling* 15, no. 1 (February 1, 2018), https://doi.org/10.1186/s12976-017-0074-5.

marketing companies. Welltok's Predilytics claims to be able to "reveal impactable risk at an individual level"[180] for 274 million registered individuals. Kaiser Permanente combines insurance and health care for 12 million people and aggregates data accordingly.[181] LiveRamp, the successor company to the already mentioned Acxiom,[182] partners with HealthVerity to "link patient health data and digital behavior." "Patient journey touchpoints can be connected from ad campaign impressions and brand website views to doctor visits and prescription fills."[183] Very soon, through a partnership with "the largest supermarket chain in the United States", this will extend to "linking the grocery carts of patients with their healthcare data and exploring how diet, smoking, alcohol consumption, Over-The-Counter purchases or even food insecurity really impact their journeys."[184] Stimulating health and pharma companies digital advertising remains a key goal to the company, as it currently accounts "for only 2,8% of total U.S. digital advertising expenditures".

China also beckons, having joined in 2017 an international body that sets quality specifications, and eased access to local databases for foreign pharma companies. Health big data and research projects are mushrooming. Sanofi, for example, conducts diabetes' and

[180] Welltok, "Analytic Services - Welltok - Optimizing Health, Maximizing Rewards," *Welltok,* 2019, https://www.welltok.com/analytic_services/.

[181] Kaiser Permanente, "Kaiser Permanente 2018 Annual Report," *Kaiserpermanente.org,* 2018, https://healthy.kaiserpermanente.org/static/health/annual_reports/kp_annualreport_2018/?kp_shortcut_referrer=kp.org/annualreport.

[182] See Introduction, page 10.

[183] HealthVerity, "HealthVerity and LiveRamp Develop Privacy-Centric Linkage between Patient Healthcare Data and Digital Behavior," *Prnewswire.com,* October 15, 2019, https://www.prnewswire.com/news-releases/healthverity-and-liveramp-develop-privacy-centric-linkage-between-patient-healthcare-data-and-digital-behavior-300938656.html.

[184] HealthVerity, "Grocery Data: The Missing Ingredient in The Patient Journey," *Healthverity.com,* October 15, 2019, https://info.healthverity.com/healthverity-8451-webinar.

immunological diseases' tests in Chengdu.[185] Delocalization of health data sources is no panacea however. Vital and other health data differ across populations. For commercial purposes, obtaining market authorization for a new drug cannot be based on tests performed elsewhere on a different population. For this medical reason alone, and for reasons of international competition among pharma firms and health providers, including analytical tools, France should keep working at facilitating access and use of its large health data bases, while guarding itself from the excesses noted above.

India's Demanding Legislation Under Preparation

Currently, the legal framework covering digital health data is simply a reference within the personal data section of the Information Technology Act (2000), mandating the protection of sensitive data and preventing unlawful disclosure. But the provision only applies to "body corporates", which do not include public hospitals. The other available rule is a 2016 Electronic Health Record Standards (EHRS)[186] released by the Ministry of Health and Family Welfare (MoHFW). It lays down technical, administrative and physical standards for data collection and storage. The scope of coverage is unclear, as are the timelines for accessing patient records. Unique identification information such as URLs and IP addresses are not listed as sensitive information. The EHRS is more about standardizing digital records than about data protection. A code of ethics for doctors

[185] Takada Noriyuki, "China's Big Data Draws Big Pharma," *Nikkei Asian Review,* August 1, 2019, https://asia.nikkei.com/Business/Pharmaceuticals/China-s-big-data-draws-Big-Pharma2.

[186] The Ministry of Health and Family Welfare, India, "Electronic Health Record (EHR) Standards Version 2016 for India" (2016), https://mohfw.gov.in/sites/default/files/17739294021483341357.pdf.

remains vague in its prescriptions. There are no laws in India mandating hospitals to disclose security breaches. By contrast, the American Health Insurance Portability and Accountability Act (1996)[187] requires that a hospital discloses a breach that has affected more than 500 patients. The GDPR also has strong provisions in case of breaches.[188] The lack of a proper regulation is also highlighted within the controversy over the extended use of Aadhaar's unique identification number and its vulnerability.

This situation will very likely undergo major changes. On the one hand, the National Health Policy (2017) includes ambitious plans for the digitalization and national integration of health data, including national health registries, platform and exchange networks, optical fiber connections and the general use of tablets and smartphones. Apps in this area are burgeoning. On the other hand, India is about to pass a draft Digital Information Security in Healthcare Act (DISHA), proposed by the Ministry of Health on March 11, 2018.[189] The period for stakeholder comment ended on April 21, 2019, and a bill is currently being finalized: even if the government is now committed, the process with the Lok Sabha (Lower House of the Parliament) could still change the outcome, as it does for many legislative acts. As of now, DISHA is a radical and all-encompassing

[187] Office for Civil Rights, "Breach Notification Rule," *HHS.gov,* September 14, 2009, https://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html.

[188] Akhil Deo, "Without Data Security and Privacy Laws, Medical Records in India Are Highly Vulnerable," *The Wire,* January 27, 2017, https://thewire.in/law/without-data-security-and-privacy-laws-medical-records-in-india-are-highly-vulnerable.

[189] Government of India, Ministry of Health & Family welfare, "Government of Lndia Ministry of Health & Family Welfare (EHealth Section)" (2018), https://www.nhp.gov.in/NHPfiles/R_4179_1521627488625_0.pdf.

proposition[190] for the protection of health data privacy, going much farther than the Modi government's other major digital law under preparation, the Personal Data Protection Bill (PDPB).[191] Individual consent is paramount, with exceptions specified much more narrowly than under the proposed PDPB in general – or the GDPR for that matter. Denial of service is impossible. The Act is also stronger in asserting the right to erasure. Government access to health data is restricted to strict health purposes. "Insurance companies shall not insist on accessing the digital health data of persons who seek to purchase health insurance policies or during the processing of any insurance claim": this is qualified only by user consent for access to the digital data held by the specific clinical establishment to which the claim relates" (Article 29.5). Pharma companies have no access to individual digital health data, even for research. A National Electronic Health Authority is to be set up, and the Act now includes provisions for sanctions in case of breaches.

The DISHA draft can still collide with the coming conclusions of the Srikrishna Committee on personal data protection, and there is professional criticism of its most radical dispositions.[192] The regulatory effort does not stop there. The Union government is currently in its last stage (with the Upper House) of a DNA Technology (Use and

[190] Singh Madhur, "India to Be First to Protect Health Data of Citizens with Iron-Clad Law?," *Business Standard,* May 31, 2018, https://www.business-standard.com/article/economy-policy/india-to-be-first-to-protect-health-data-of-citizens-with-iron-clad-law-118053100126_1.html.

[191] Ikigai Law, "DISHA and the Draft Personal Data Protection Bill, 2018: Looking at the Future of Governance of Health Data in India," *Ikigai Law,* February 25, 2019, https://www.ikigailaw.com/disha-and-the-draft-personal-data-protection-bill-2018-looking-at-the-future-of-governance-of-health-data-in-india/#acceptLicense.

[192] For an example of these criticisms, see: Rahul Matthan, "A New Direction for Data Privacy in Healthcare," *Livemint.Com,* April 11, 2018, https://www.livemint.com/Opinion/3LK0TR6zdXmeIkaJuTUnnJ/A-new-direction-for-data-privacy-in-healthcare.html.

Application) Regulation Bill 2019, regulating and limiting the use of DNA profiling to civil paternity suits and to consented processing of DNA data for all crimes that are subjected to the Indian Penal Code, of which crimes liable to more than 7 years imprisonment do not need consent.[193]

Overall, India's handling of the data protection and privacy issues in the health sector seems unique. A very strong push for coordinated and integrated digital tools coexists with what promises to be a stringent privacy policy – surpassing the GDPR's requirements in several ways. So far, while health care at the grassroots is likely to be enhanced by the drive for digitalization, pharma research – whether it is conducted by foreign or Indian companies – would seem to be the least of the Indian government's priorities. DISHA stands in contrast with the overall trend favoring innovation and state requirements over privacy protection, as evidenced by the PDPB and with numerous digital policies.

China's Use of Health Data As a Resource

Health data in China is explicitly viewed as a resource for the developmental state. The incitation of health data usage comes as part of the "Internet Plus" strategy proposed by Prime Minister Li Keqiang in 2015, which aims at boosting the development and economic value of some conventional industries through the use of internet. "To let the people run less, and to let the data run more", is the widely used phrase to explain the concept of "Internet Plus."

[193] Ministry of Science and Technology and Earth Sciences, "The DNA Technology (Use and Application) Regulation Bill, 2019," July 8, 2019, https://www.prsindia.org/billtrack/dna-technology-use-and-application-regulation-bill-2019.

Not surprisingly, the "Guiding Opinions on Promoting and Regulating the Application and Development of Big Data in Health and Medical Care", issued by the Chinese State Council in 2016, describe health and medical big data as a fundamental and strategic resource of the state. This occurs in the context of fostering new business sectors and enabling more economic growth.[194] It then goes on to emphasize the need for better use of the government, to provide top design for "the integration, sharing and open application of big data in health and medical care" and to "provide powerful support to the building of a healthy China, comprehensively finishing building a moderately prosperous society, and the realization of the Chinese Dream of the great rejuvenation of the Chinese nation."

With governmental support and pushes for the integration, sharing and open application of big data in health and medical care, the efforts are bearing results. Taking the example of Guangdong province, data-sharing is achieved within 3112 medical and health institutions, and the provincial-level national electronic health data bank holds information on 80 million permanent residents.[195]

One detailed presentation on China's digital data policies notes that "although privacy is an extremely important topic for big data in health and medicine, there is no specific law or guidance on this in China."[196] Indeed, the Cybersecurity Law of 2017 has no mention

[194] General Office of the State Council of the People's Republic of China, "Guiding Opinions on Promoting and Regulating the Application and Development of Big Data in Health and Medical Care 国务院办公厅关于促进和规范健康医疗大数据应用发展的指导意见," (2016), http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm.

[195] Health Commission of Guangdong Province, "Guangdong Health Case Letter," *Gd.Gov.Cn,* June 19, 2019, http://wsjkw.gd.gov.cn/zwgk_bmwj/content/post_2516919.html.

[196] Luxia Zhang et al., "Big Data and Medical Research in China," *BMJ Medical Research,* February 5, 2018, j5910, https://doi.org/10.1136/bmj.j5910.

of the word "health", and the PIS included "health information" in the definition of sensitive personal information without going further. However, three drafts of the Basic Healthcare and Health Promotion Law have been published for public comments since 2017. The draft law addresses the issue of health information privacy in the following way in Article 90 (in the third draft, for public comments until September 26, 2019): "The State protects the personal privacy related to the health of citizens and ensures the safety of personal health information. No organization or individual may acquire, use or disclose citizen's personal health information except as required by law, administrative regulations or with the consent of the person."[197]

In April 2018, the General Office of the State Council published the "Opinions of the General Office of the State Council on Promoting the Development of 'Internet plus Health Care",[198] again with no mention of the word "privacy". This was followed by the issuing of "Measures for the Administration of Internet Diagnosis and Treatment", "Measures for the Administration of Internet Hospitals" and "Specifications for the Administration of Remote Medical Services" in September 2018, all three for trial implementation.[199] The latter three do have a general mention of "privacy protection", without going much into the actual means for achieving this. Nevertheless,

[197] National People's Congress, "Basic Healthcare and Health Promotion Law (draft) 中华人民共和国基本医疗卫生与健康促进法(草案)," (2019), https://npcobserver.files.wordpress.com/2019/08/basic-healthcare-and-health-promotion-law-3rd-draft.pdf

[198] General Office of the State Council of the People's Republic of China "Opinions of the General Office of the State Council on Promoting the Development of 'Internet plus Health Care 国务院办公厅关于促进"互联网+医疗健康"发展的意见" (2019)

[199] The National Health Commission of the People's Republic of China, "About the issuing of the Measures for the Administration of Internet Diagnosis and Treatment(trial implementation), etc. 关于印发互联网诊疗管理办法（试行）等3个文件的通知," (2018), http://www.cac.gov.cn/2018-09/14/c_1123431844.htm

they all take on the issue of cooperation with third-party institution, and stress the need of an agreement specifying the responsibilities of all parties in various areas, including on privacy protection. Hence, the statement in the above-detailed presentation is technically inaccurate, but the authors' comment probably reflects the distance between law and practice.

So far, the most concrete regulation on health data protection is the "Administrative Measures on the Standards, Security and Service of National Health and Medical Big Data (For Trial Implementation)"[200] issued in July 2018 by the National Health Commission. It tackles the issue of data collection, data storage, service provision, data utilization and data sharing.[201] Specifically in terms of data sharing, "the National Health Commission is responsible for establishing an open sharing mechanism for healthcare big data, coordinating the construction of a resource catalogue system and a data-sharing exchange system, and strengthening the service and management of the health care big data life cycle." The goal is tilted towards the construction of shared health data resources but the rules do little to specify health data protection beyond the generalities already present in the PIS.

[200] The National Health Commission of the People's Republic of China, "Administrative Measures on the Standards, Security and Service of National Health and Medical Big Data (For Trial Implementation) 关于印发国家健康医疗大数据标准、安全和服务管理办法（试行）的通知," (2018), http://www.cac.gov.cn/2018-09/15/c_1123432498.htm

[201] The National Health and Family Planning Commission of the People's Republic of China, "Explaining the Administrative Measures on the Standards, Security and Service of National Health and Medical Big Data (For Trial Implementation) 国家健康医疗大数据标准、安全和服务管理办法（试行）》解读", September 14, 2018, http://www.cbdio.com/BigData/2018-09/14/content_5834771.htm

China has gone all the way in the direction of collecting, aggregating and using all available health data for health care and the pharma industry. India is contemplating a very restrictive law, although final developments deserve to be watched. The interdependence of private insurance and health care in the United States is troubling, but this is also where the developments from AI and big data are the more promising: this is a battle ground in the UK, due to the large base of its National Health Service. The French case should serve as a reminder that usable big data is not so easy to gather from decentralized systems. Provided the rules are set clearly on what cannot be used or even accessed by insurance companies - including with consent by individuals - a key goal for public health should be to improve the range, accessibility and quality of medical, genetic and behavioral data.

VII

CHASING PRIVACY, INNOVATION AND PUBLIC INTEREST

Regarding processes and methods for personal data protection, the contrast between the European and American approaches exists, but it should not be overstated. The GDPR may look like an orderly French garden, and the U.S. regulations as a maze, but we should go beyond the appearance of texts. The *"Qui veut trop embrasser mal étreint"*[202] adage could still be invoked against the GDPR and many other EU directives. Sundar Pichai, Google's CEO, warns against a general approach and advocates for a sector-by-sector regulation on AI, "rather than rushing into a way that prevents innovation and research."[203]

Appeals to the Court of Justice of the European Union (CJEU) will create precedents and to some extent work as case law does in the United States. Recent examples, such as an October 1st, 2019 ruling by the CJEU on specific rules for user consent to cookies, indicate that this is happening.[204]

On exceptions related to public interest, the distance between the United States and Europe is also likely to decrease – and this might be unwelcome news. We have already emphasized the similarities of data scraping between market analysis (or "surveillance capitalism"

[202] "Grab all, lose all".

[203] Tim Bradshaw, "Google Chief Sundar Pichai Warns against Rushing into AI Regulation," *Financial Times,* September 20, 2019, https://www.ft.com/content/b16e6ee8-dbb2-11e9-8f9b-77216ebe1f17.

[204] CJEU, "Bundesverband der Verbraucherzentralen und Verbraucherverbände–Verbraucherzentrale Bundesverband eV v" Planet49 GmbH (October 1, 2019).

as some will call it) and state surveillance. But the potential for abuse of public interest exceptions, whether through biased interpretation of the law or by illegal practices, is always there. Edward Snowden's chilling description[205] of the mass surveillance programs created after 9/11 and how they went around constitutional protections should be studied, whether one thinks he is a whistleblower or a traitor. Polemics around biometric identification are not reserved to India's Aadhaar program. France has just decreed a one-time facial recognition process to verify the identity of passport holders and foreigners applying for residency who contact public services.[206] As Europeans catch up with big data analytics and AI programs, the issues will continue to grow.

Looking at what partners and competitors do – who may occasionally go beyond the GDPR or more often undercut it – and ways to improve and revise it, we shall make some policy suggestions, both positive and negative. These are highlighted below.

Ambitious but Generic Rules

Explicitly, the GDPR is a "general regulation", a term previously not employed for EU regulations. This begs the question of explanatory guidelines and sectoral regulations. There has been a few of the former, none of the latter. **Ambitious but generic rules leave room**

[205] Edward Snowden, *Permanent Record* (New York: Macmillan, 2019).

[206] For a criticism by the French Supervisory Authority, see
Source: CNIL, "Délibération N° 2018-342 portant avis sur un projet de décret autorisant la création d'un traitement automatisé permettant d'authentifier une identité numérique par voie électronique dénommé "Application de lecture de l'identité d'un citoyen en mobilité" (ALICEM) et modifiant Le Code de l'entrée et du Séjour des étrangers et du droit d'asile (demande d'avis N° 18008244)" (October 18, 2018).

for interpretation and loopholes, and create large spaces for exceptions. We come to the paradox of Article 23, soberly entitled "limitations." It provides for state access to citizens' data with a suspension of state obligations, as well as citizens' rights, in certain cases.

Above all, **the stronger and broader the requirements for compliance, the more non-implementation is likely.** We all meet with examples of glaring non-compliance – simply starting with websites where it is impossible in practice to make any choice about one's data privacy. Digital companies, and representatives from the business world, which are now controllers or processors, point out to several factors: the cost of compliance, especially for smaller companies; the issue of human resources – trained data officers, for example. They regularly stress that the very language of the GDPR emphasizes very different objectives, which are difficult to reconcile. **There is a lack of operational help to companies, and a lot of guidance is needed to make them compliant.** A large responsibility for interpretation is vested with the controllers and processors.

The Fault with Opt-Out by Default

The "notice and consent" approach is perhaps the most popular aspect of the GDPR because it is said to put the individual back in control of his personal data. And more control seems more satisfying. In reality, **a maximal approach has its drawbacks, and it seems a nuclear option.**

A negative default option (unless a positive or "opt-in" action is taken by the user, no data or metadata can be lifted) or an overall

do-not-track option would greatly reduce the personalized use of most websites and apps, which are based on data queries and exchanges. A GDPR requirement is that access cannot be conditional on consent on data extraction. An individual's rational choice is indeed to get the product while refusing to turn over the information. But if many users expect to free-ride the web at the expense of those other users who still opt-in, this is likely to kill the economy of the internet, which has been based on free availability against personal data. It is also just as unfair as the present situation, where very unequal compliance from websites and companies gives a competitive advantage to sinners over those who implement the letter of the GDPR. **At the macro level, this bankrupts the system and ends with the same result as the full negative default option. Someone has to pay, one way or another, or we will be back to the pre-internet age. Would the internet be the same if all services and information worked on a paying basis?** It is a fundamental change that may not be accepted by individuals as consumers, unlike citizens claiming privacy rights. A French liberal economics think tank has proposed to reverse the proposition: individuals being the owners of their personal data could sell it to the digital platforms. Obviously the proposal will immediately run into many issues – third party or unpredicted use for example. But it has the merit of ending the hypocrisy over the relationship between internet users and providers.[207]

[207] Isabelle Landreau et al., "Mes Data Sont à Moi - Pour Une Patrimonialité Des Données Personnelles," *Génération Libre,* January 2018, https://www.generationlibre.eu/wp-content/uploads/2018/01/2018-01-generationlibre-patrimonialite-des-donnees.pdf.

The Privacy Paradox

Against the unsustainable consequences of a rational choice by individuals, we should consider what privacy experts call the "privacy paradox"[208] but that could equally be termed the "hypocrisy of trust". Need, desire and lust – not necessarily in that order – override the precautionary principle in the individual psyche. In a digital environment, we constantly surrender our privacy to other useful or pleasing purposes. What one could call "speed googling", certainly runs against reading privacy notices. This is only a tiny example in a world full of trade-offs between convenience and data protection.

The coming disappearance of cash (already a reality in countries as diverse as China and Sweden) is a huge case in point, leading to hypocrisy of trust. China was once the world's most cash-loving society, in part because of its anonymity, in part because it is the instrument of symbolic exchange among individuals – including to departed ones. Today, even the most minute cash transaction is likely to be replaced by electronic payments. **A cashless society is also a transparent society, completely reversing the anonymity that cash brought to transactions.** Hence, after the one-hundred-dollar bill, cryptocurrency, such as bitcoins, became the new anonymous currency of choice, so long as it is a private medium of exchange. China then decided to become the world's first public emitter of cryptocurrency of a digital currency, which can no longer be called a cryptocurrency. This could conceivably help China evade the dollar's grasp on international transactions. But the Bank has already announced that it will "tag" the currency to trace the individuals who used it. "The Central bank digital currency can be

[208] See II. What Is Privacy and How Can it Be Ensured?, *Privacy Policies, Notice and Consent,* page 43.

circulated as easily as cash (…) at the same time, it can achieve controllable anonymity", according to Mu Changchun, Deputy Chief in the Payment and Settlement Division of the People's Bank of China (PBOC)."[209]

Differences Between Member States

Differences of interpretation are also reflected by the various adaptations into the national law of member states.[210] **Age of consent is most often mentioned as an issue. There are also huge differences in enforcement capacities among member states,** and probably different degrees of willingness to move ahead. A study including 17 EU member states as well as Croatia (but without countries such as France, the Netherlands or the UK) finds highest level of personnel resource (circa. 250) for Polish and Spanish supervisory authorities and less than 50 people for each of the other12 countries.[211]

Since national security is not a competency of the EU, member states can specify for themselves what constitutes (or what doesn't), further widening the scope of derogations. *Quis custodiet ipsos custodes?*[212]

[209] Dexin Guo, "'Digital Renminbi' Is Revealed '数字人民币' 初露真容," *Xinhua,* August 21, 2019, http://www.xinhuanet.com/fortune/2019-08/21/c_1124900323.htm.

[210] For a review of national adaptations after eight months, see this ten member states case study:

Source: Karen Mc Cullagh, Olivia Tambou, and Sam Bourton, *National Adaptations of the GDPR* (Luxemburg: Collection Open Access Book, Blogdroiteuropeen, February 2019).

[211] EDPB, "First Overview on the Implementation of the GDPR and the Roles and Means of the National Supervisory Authorities," March 8, 2019, https://edpb.europa.eu/sites/edpb/files/files/file1/19_2019_edpb_written_report_to_libe_en.pdf.

[212] "Who will guard the guardians?"

Decisions on interpretation of the GDPR can follow different processes: if the issue involved has cross-border implications and if the individual or entity raising it has a seat inside an EU member state, that member state's supervisory authority will take the leading role in examining the case, and its decision will be valid across the EU: that is the one-stop shop approach. **But the process does not apply if the plaintiff is not based inside the EU, or if the locus of decision-making does not coincide with the legal establishment.** There can thus be 28 decisions. Even this has its own exception: in "urgent" cases, any supervisory authority can make a decision that will stand only for three months. One would have hoped for a simpler and clearer process.

Three recent cases, all involving Google as it happens, will serve as illustrations of the complexities involved. On the issue of Google's staff recording private conversations in order to improve voice-recognition performance, the Hamburg Commissioner for Data Protection was able to take measures, even though Google's main establishment is in Ireland.[213] A landmark decision by the CJEU establishes that Union law (the GDPR) does not mandate the implementation of a de-referencing obligation in search engines beyond the EU's borders.[214] But a member state jurisdiction can indeed impose de-referencing on all versions of a search engine "in light of national standards" and by examining the balance between the right to privacy and freedom of information. In other words, national jurisdictions can go beyond, but not undercut the GDPR in

[213] The Hamburg Commissioner for Data Protection and Freedom of Information, "Press Release. Speech Assistance Systems Put to the Test - Data Protection Authority Opens Administrative Proceedings against Google," August 1, 2019, https://datenschutz-hamburg.de/assets/pdf/2019-08-01_press-release-Google_Assistant.pdf.

[214] CJEU. Google LLC, successor in law to Google Inc. v *Commission nationale de l'informatique et des libertés* (CNIL) (September 24, 2019).

their own legal practice, provided a "balance" of different rights has been examined. In another case waiting for a CJEU judgment, the French CNIL has decided that, if a company's decision-making in GDPR makes it a "controller" –, and it is located in a different member state than the one where the company has its establishment, the case can be decided by that other country's supervisory authority. Given the difficulty to locate activity and decision-making, identifying the actual locus of company decisions will prove contentious, and in any case defeats the simplicity of the one-stop shop provision. The decision "may ultimately prove detrimental to effective EU-wide enforcement (including uniformity in application and legal certainty) in the longer term."[215]

That Google – with an EU establishment in Ireland but much larger interests in other member states – has often been a test case is no accident. There is a trend towards more sovereignty over digital data as well as over fiscal resources, and therefore towards the determination of the location according to the market. However, this trend is contested and indecision will harm the prospect of a unified digital market, and any company that manages data across borders in general. **A radical shift, whether for regulation or for taxation, from location of establishment, production to point of sale or place of decision, is a large-scale international undertaking.** Simultaneously mixing the two approaches is like having some vehicles drive on the right side of the road while having others drive left…

[215] Lokke Moerel, "CNIL's Decision Fining Google Violates One-Stop-Shop," *SSRN*, February 19, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3337478.

Technology Is a Fast-Moving Frontier

Technology is a fast-moving frontier where existing rules cannot anticipate, except in very general terms and objectives. The train cannot be stopped, unless one decides a radical check on innovation. Would our competitors consider implementing this? Hardly, if one judges from the Chinese, Indian and American cases. The issues go beyond the "Geneva Convention", which some have called for. For privacy and data protection, as with cybersecurity, and in fact arms control negotiations in general, international agreements are welcomed – but they require robust verification. In cybersecurity as in arms control, deterrence has often emerged as the complementary option or the alternative to agreements. A similar option does not exist for privacy rights. **It is only by opening or closing our digital market that we can hope to influence the behavior of actors whose base is outside of it.**

We must know that there are unknown unknowns. It is impossible to predict which type of data will turn out to be personal, sensitive or critical. The situation for AI is reminiscent of a recent advertising billboard from a savings company: "robots can't take your job if you're already retired":[216] by implication, everybody else is up for grabs. Some examples follow, although they pertain to cases that have already materialized by definition: DNA data banks, such as 23andme or Ancestor.com, are revolutionizing criminal enquiries by re-identifying anonymous DNA from distant relatives (third- or fourth-degree cousins), as was the case in 2018 for the notorious Golden State killer in California. A 2 million people genetic data bank is

[216] Robots, of course, are an outcome of AI. This 2019 Prudential billboard has justifiably received its own publicity on the web. Among numerous websites :
Source: r/ABoringDystopia, "Automation Can't Take Your Job If You Don't Have One.," *Reddit,* January 29, 2016, https://www.reddit.com/r/ABoringDystopia/comments/bci7pz/automation_cant_take_your_job_if_you_dont_have_one/.

enough to identify 90% of the American population.[217] Lenddo, a Singapore-based technology company, uses online social and mobile behavior (such as whether you keep your smartphone battery filled up) to determine the borrower's "willingness to pay" back loans.[218] Platforms for streaming TV channels now track viewers and monetize them through behavioral advertising.[219] Social network data analysis, helped by integration and fusion techniques, can pinpoint probable characteristics of individuals far better than any technique previously available.

Blockchain technologies create an additional problem for one aspect of privacy – the right to erasure. In blockchains, all participants can view all data recorded; several copies of the blockchain coexist on different computers; once data is recorded, it cannot be altered or removed; and decisions are made by consensus between participants, without a central arbitrator. Precautionary steps such as encryption and pseudonymization are definitely required.[220] But even with these, the French CNIL and the UK's Financial Conduct Authority (FCA) warn at present against the use of blockchains.[221]

[217] Yaniv Erlich et al., "Identity Inference of Genomic Data Using Long-Range Familial Searches," *Science* 362, no. 6415, October 11, 2018: 690–94, https://doi.org/10.1126/science.aau4832.

[218] Hope King, "This Startup Uses Battery Life to Determine Credit Scores," *CNNMoney,* August 24, 2016, https://money.cnn.com/2016/08/24/technology/lenddo-smartphone-battery-loan/index.html.

[219] Hooman Mohajeri Moghaddam et al., "Watching You Watch: The Tracking Ecosystem of Over-the-Top TV Streaming Devices," *Freedom-to-tinker.com*, September 18, 2019, https://freedom-to-tinker.com/2019/09/18/watching-you-watch-the-tracking-ecosystem-of-over-the-top-tv-streaming-devices/.

[220] CNIL, "Blockchain and the GDPR: Solutions for a Responsible Use of the Blockchain in the Context of Personal Data," *Cnil.fr,* November 6, 2018, https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data.

[221] Andrew Solomon, "Block Chain: Is the GDPR out of Date Already?," *Lexology.com,* August 30, 2017, https://www.lexology.com/library/detail.aspx?g=d4c0481a-c678-4748-80cb-4ab917e66207.

Collection versus Usage

What can more realistically be regulated is the usage of collected data and its interpretation, provided there is indeed rule of law, including a legal right of verification and appeal by individuals. In a sense, this has always existed in legal processes: an individual cannot be convicted in court on the basis of illegally collected evidence. The first barrier against the indiscriminate use of algorithms is the prohibition of individual decisions based solely on automated processing alone. Article 22 of the GDPR has in fact laid down this prohibition, while laying down very broad exceptions. Again, member state law can introduce additional safeguards. A good example in the French case has been the change of the admission process to higher education from a fully automated algorithm (Admission Post-Bac or APB) to one requiring human intervention and some possibilities for requesting an explanation regarding the decisions made (Parcoursup). **One should edict a similar limitation when it comes to autonomous driving (or flying): ultimately the principle of torts requires that human action and responsibility can be identified.** This is one of the answers to a question posed by a recent report on AI: "Are there areas where human judgement, fallible though it is, must not be replaced by a machine?"[222]

The second barrier is about **requiring checks and balances in the implementation of adverse decisions based on digital evidence.** Legislation adopted under the spell of a terrorist attack will often lower this barrier. The Lashkar-e-Taiba attacks in Mumbai (2008) or the terrorist actions in France (2015) have in both cases led to new anti-terrorist actions that offer less or no opportunity for a

[222] Cédric Villani, "Executive Summary. For a Meaningful Artificial Intelligence," March 2018, https://www.aiforhumanity.fr/pdfs/MissionVillani_Summary_ENG.pdf.

judiciary check. In the French case, they include "broad powers to search computers as well as the ability to block websites that allegedly glorified terrorism, all without prior judicial authorization."[223] A new intelligence law adopted in July 2015 is now awaiting a decision by the European Human Rights Court on its legality.[224] **A very hard look should be given at all exceptions that circumvent the need for a judicial review** or base negative action on predictive criteria rather than actual evidence. In more innocuous, but economically significant cases such as insurance, **criteria are needed to limit the individualization of customers, balancing the need to dis-incentivize risky behavior with the collective nature of insurance.**

No privacy in your own car

Autonomous driving has issues that go beyond privacy – security is one. No one would treat black boxes on planes as an invasion of the pilots' privacy, and no one would question the existence of drive record disks on commercial trucks. Drivers routinely install dash cameras that are supposed to capture the circumstances of an accident and help establish responsibilities – or ensure lower insurance rates. Other innovations do limit privacy: GPS tracking, driving apps, toll collection. As it improves, autonomous driving will create enormous issues relating to automated decisions and responsibilities. These are distinct from the privacy and data protection issue – although hacking is a serious concern for both security and privacy.

[223] David Sullivan, "The Consequences of Legislating Cyberlaw After Terrorist Attacks," *Just Security,* April 9, 2019, https://www.justsecurity.org/63560/the-consequences-of-legislating-cyberlaw-after-terrorist-attacks/.

[224] Assemblée nationale et Sénat, "Loi N° 2015-912 du 24 Juillet 2015 Relative Au Renseignement (1)," https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000030931899&categorieLien=id

But further developments, some of them hard to foresee, will create new types of privacy issues, and the systems enabling automated driving can differ widely in that perspective. In simple terms, there are two opposite methods: one is a robot-like car, enabled with a multiplicity of sensors that finds its way and avoids obstacles. This is the American path, one informed by the notion of free individuals. The dependency on GPS-like devices is no greater than with the previous generation of cars. Still one might find, that cars become computers, as is the case for maintenance. They store a record of all their previous activity, to be likely left in place at the time of resale (a Tesla stores all driving actions since being first put on the road).

The other approach is to treat the car as a smartphone on wheels, a receiving and emitting device with a network system. At a minimum, this means digital highways, and China is already pushing in that direction. Better management of traffic flows, from Waze-like applications for example, also follows that path. This choice increases on the spot control of course. But it can go further: already, breath sensors can stop an inebriated driver from starting his engine. Recognizing a pedestrian's disability or his/her proneness to drunkenness or jaywalking will improve if that personal information is already in the system. The first generation of autonomous devices is largely about passive defense. The second generation could be more pro-active and inquisitive. Again, the trade-off between privacy and security reappears.

With this leap, a car becomes a major collection of devices in the IoT. Data is shared among many parties. **Manufacturers, insurers, traffic managers and car sharing providers may be combined controllers when they jointly determine the means and purposes of processing certain personal data.**

Sovereignty and the Splinternet

Data sovereignty has a difficult relationship with data privacy. States attempting to move towards data sovereignty actually want more control over what they call *their* data. This often coincides with a disdain for digital privacy, and a fight against instruments – clouds out of their reach, encryption and messaging apps, VPNs and the like – that may shield any individual's personal data and communications. One should also honestly recognize that the frontier is porous with the legitimate requirements of any rule of law society, and that this is therefore in a continuum with discussions on public order, national security and the "general interest." Attitudes can also differ across different categories of data: we have seen, for example, a prospective health data bill in India being far more protective of privacy than American laws or the overall European framework. Different attitudes towards data sovereignty and digital privacy lead to the fragmentation of the internet – creating a "splinternet." This could either take the form of like-minded groupings or follow national divisions based on digital sovereignty.

The first trend may be achievable for like-minded states that subscribe to common values, to the rule of law and above all, to a degree of mutual oversight. The European GDPR, and the ensuing adequacy agreements with third countries, are as much predicated on the free flow of data as they are on data or privacy protection. Japan has also launched a "Data Free Flow with Trust" (DFFT) initiative at the June 2019 Osaka G20 summit. The initiative aims at shaping global data governance, but is even more aimed at warding off data sovereignty (and in particular China's Great Firewall model…) than at regulating data privacy: it has been met with limited success so far, India, Indonesia and South Africa refusing for example to sign

a joint declaration. OECD Guidelines on the protection of personal data and on trans-border flows (1980, revised in 2013) point to the necessity of a gradual convergence. The Asia-Pacific Economic Cooperation (APEC) also issued a Digital Privacy Framework (2005, revised in 2015) patterned on the OECD guidelines.

The second approach – prioritizing digital sovereignty and splintering the global internet – is much more likely in cases where the state prevails over the law, and where mistrust of other states is the norm. For example, Russia and China may look very much alike in their concepts of sovereignty, surveillance, and cyberactivism abroad. But precisely for these reasons, how could they constitute a joint data environment with free flow? Just as in the long run for market economies vs. state-driven policies, **societies implementing the rule of law may have the advantage, because they are better equipped to exchange big data with some degree of security. However, they will reflect on the wisdom of allowing platforms from closed environments such as China's or Russia's to poach on their open data markets.**

Towards the Dual Internet World

A dual internet is the next best alternative to the single WorldWideWeb. Assuming that some fragmentation of the digital world is unavoidable – and in fact wished for by the state exponents of a closed internet and data sphere – personal data and privacy protection requirements favor a two-world solution, where one implements domestic and cross-border rules, while the other could fragment according to national and state control boundaries. In the real world, the choices are not as clear-cut: different states will have different requirements

on some categories of data, and trust cannot be at the expense of verification, even with the closest partners. Let us therefore recognize that **adequacy "decisions" for free data flows will need to factor in a strategic goal of avoiding isolation.**

Global issues beyond the EU's borders should be anticipated from the bottom when considering the rules and the implementation of personal data and privacy protection. There are numerous reasons for this. One rests within the sovereignty and digital fragmentation debates. A second reason is that the very concepts and experiences in this field are both evolving and shared. Even if they do have an overall consistent philosophy as a starting point in principle, Europeans tend to exaggerate how much they are at the source of thinking on these issues. Time and again, we encounter cases and debates, legislative examples and above all data protection techniques (and challenges to these techniques) that originate in the United States. It is not perceived enough that, in spite of the strong normative influence that the GDPR exercises, **many solutions and tools for data protection originate in the United States.**

It is beyond the reach of the present study to discuss the financial and institutional implications of these technological choices for European innovation. Europe has in fact some of the talent, often captured by the companies that possess a competitive edge. But it is clear that searching for optimal solutions requires **an open mind and cooperation over the Atlantic, and in fact with all broadly like-minded partners confronted with the same issues.** Above all, the rule of law is the common denominator, above and beyond the close data integration that exists among these partners. And we should not assume that the **rule of law is ensured without verification and independent institutions – in other words, without checks and balances.**

Our study of India and China also indicates that a global fight among models of digital governance may well be under way, as is the case with other global issues. The triangle formed by privacy, efficiency and security offers different solutions. As is the case with trade and financial flows, Europe cannot be conceived as a closed digital universe. Beyond transatlantic cooperation and compromises over these issues, **the will to create norms should be balanced with the need to remain attractive.** An analogy can be made between the choices for Europe's "adequacy decisions" (in fact, adequacy agreements) and trade agreements. The latter comes in different varieties, from surface to deep trade agreements. The trend of the last decade had been to build more and more comprehensive treaties incorporating investment and arbitration, services, intellectual property and norms:[225] until the trend backfired with the melt-down of the Transatlantic Trade and Investment Partnership and initial Trans-Pacific Partnership projects. **Adequacy decisions will have to make similar choices, between demanding and comprehensive criteria requiring strong and permanent adjustments by partners, and more limited data-sharing agreements** (of which, even if controversial, the Privacy Shield is the best example).

Ex-Ante or Ex-Post Action

Neither the notice and consent framework, the trust and privacy by design approach nor the ex-ante regulation can be rejected on the grounds that they are limited. They do serve a purpose towards privacy and data protection. But **regulation must also rely on ex-post**

[225] Edith Laget et al., "Deep Trade Agreements and Global Value Chains," *World Bank Group,* June 2018, http://documents.worldbank.org/curated/en/356541529933295649/Deep-trade-agreements-and-global-value-chains.

action and that is largely the road towards penalties or tort litigation. The two avenues are not identical, although they can be combined: case law and tort litigation are the customary avenues in the United States, but regulatory agencies can impose very steep fines on offending companies. By contrast, the European tradition lies within explicit provisions for penalties rather than through litigation procedures.

In both cases, numbers matter. If co-operative regulation gives way to responses after the fact, sanctions must be graded – steep penalties serve as a deterrent to the largest or more egregious offenders. The main issue with this approach is that, while digital technology scales – and therefore the size and number of offenses can be gigantic, given the public affected –, **regulatory agencies, courts and law enforcement authorities do not scale.** And since many of the offences are not perceived by the general public as much more than a pinprick, if at all, the number of complaints does not reflect the magnitude of the problem. Class actions are especially needed in this area. It should also be known if data has been illegally collected, to what use it may have been put, and what is the connection between harm suffered and the use of that data by another party. These considerations are important, because penalties can either be proportional to a crime on the book, or commensurate with the actual amount of harm inflicted. This is more difficult to assess in digital cases.

In the United States, the Federal Trade Commission (FTC) has a very big stick, and is very seldom contested in courts by the targeted companies. But it has used this stick sparingly in the past, with a graduated approach in privacy cases. It often goes for negotiated settlements, or “consent decrees” with digital technology firms such

as Facebook, Google, Microsoft, Twitter, Snapchat, and Oracle.[226] This has the advantage of putting the company involved under the equivalent of probation for as much as 20 years. Repeat offenses can have larger consequences – as much as 16,000 USD per individual offense, multiplied by thousands or millions of cases. Until recently, penalties for breaches of privacy rules had never reached the level of fines imposed on anti-competitive practices. Google, fined first in 2011, and again in 2012 under the same offense, had to pay 22,5 million USD the second time around. The amount was considered large at the time. It is now dwarfed by the 5 billion USD fine imposed on July 24, 2019 on Facebook (or 9 % of Facebook's turnover in 2018) for deceptive privacy practices, which were partly tied to the Cambridge Analytica case. Evidently, there is a change of scale that brings privacy issues level with that of competition-linked cases.

In Europe, only one fine (again, CNIL v. Google) surpassed the million-euro mark in the first year. There was another case in Denmark since; but quite ironically, **it is the departing United Kingdom that has changed the scale, with two fines of respectively 99 and 283 million EUR, in both cases for breaches affecting a very large number of individuals.**

The alternative to sanction is torts litigation, or what one inspired expert calls the "internet of torts." The model is evidently based on the U.S. consumer-based approach to privacy, and is being increasingly applied to data security breaches originating from a data fiduciary – in other words, the guardian or operator of the data. The reasoning is simple: **a data fiduciary (or any company) should take**

[226] William McGeveran, "Friending the Privacy Regulators," *Arizona Law Review* 58, no.4 2016: 959–1026, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2820683.

the necessary precautionary measures if their cost is not higher than the damage from a breach, weighed by the probability of the damage.[227] The model is the so-called Hand Formula: it was first used by a judge for an assessment of damages and liability after the sinking of a barge in 1943 where the shipping company was found guilty.[228] It proceeds fromore an economic recognition – companies do cost/price analysis. In practice, of course, the formula is harder to implement, and even more so in assessing the virtual, moral or reputational damage that can flow from a privacy breach.

Conclusion

To a large degree, much of our privacy has vanished forever, or as long as the digital age lasts. Since AI and sophisticated algorithms are not –(yet)–available to your next-door neighbor, it still makes sense to engage in regulations covering the collection and processing of personal data. The larger and more sophisticated entities – whether they are platforms, digital companies or well-endowed states – are very likely to defeat some of these rules at least some of the time, either through restrictive interpretations, through swamping the capacities of implementation, or simply because existing rules fail to cover new categories of collected data and their interpretation.

[227] Rebecca Crootof, "The Internet of Torts," *Duke Law Journal* 69, February 26, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3342499.

[228] The U.S. vs Carroll Towing Co. Case is apparently taught in every law class; United States v. Carroll Towing Co., 159 F.2d 169 (2d Cir. 1947), Justia Law (US Court of Appeals for the Second Circuit 1949), https://law.justia.com/cases/federal/appellate-courts/F2/159/169/1565896/

The above are only hints of some of the practical and ethical problems that we will face in the digital and AI age. Our tour d'horizon has focused on existing or planned regulation, on the present or near future, and it already shows glaring gaps. Almost all of the rules established so far, from top-down China to cartesian Europe, not to mention India's mix, seem to address the first basic issues of this age, such as protecting the state or public interest, ensuring user awareness and consent or regulating the processing of e-data. Again, innovation and technology run faster.

What follows is therefore **a short inventory of recommendations designed to improve these rules.** Charting the path for AI and thinking of the mixture of encouragement and guardrails that AI will require in order to remain relatively harmless, is a work of a different scale. It requires first of all, a contribution from IT scientists who are aware of the potentialities of AI and its modes of operation. We will be content with writing some recommendations for the present age as they emerge from the field that we have been able to survey.

1. Strengthening the GDPR's oversight, enforcement and adaptability

The first proposition is a familiar one for any EU regulation – a rule can only be as good as its actual implementation.

With 28 national supervisory authorities possessing widely different levels of resources, this is not an easy proposition. The conformity of national decisions with the GDPR is likely to be tested by appeals, such as the one currently in front of the CJEU regarding a defamation law.[229] Throughout these appeals, the one-stop shop's issue will also be tested – because supervisory authorities claim their own roles, and because there is a suspicion that the weaker regulatory instances will be laxer, and therefore more often retained as digital companies' center of operation.

As it is currently doing in several instances, the CJEU recognizes that the GDPR does not empower one-stop shop in many cases, and that it distinguishes formal establishment from actual decision-making and control. The CJEU decision is by definition legally sound about the existing GDPR. But from this flows a proposition: **a revised GDPR should prevent restrictions that damage a unitary decision-process.** This process is key to a single digital market in the future. This also has material implications: **guidelines should be set for**

[229] In this case, a leader of Austria's Green Party, targeted by defamatory (according to Austrian law) posts on Facebook, is asking the CJEU to mandate worldwide erasure of this and any "similar" posts. The case also involves the issue of automated treatment.

budgetary and human resources allocated by each member state to their supervisory authorities, taking into account the size of the country but also the density of digital actors. Interestingly, Ireland, which is the most obvious focus of digital companies beyond what its size would warrant, has actually taken the lead both in enforcing the GDPR and in ruling on cases, so that it would not be greatly challenged by a requirement for additional resources.

The Commission's and multi-stakeholder committee's stock-taking reports after one year both emphasize the low level of penalties imposed so far (as we have seen, it is the UK's supervisory authority that has changed the scale in recent months). It may be too early to judge, as there are still ambiguities and possible misunderstandings in the implementation of the GDPR. Guidelines are being issued: there will be a dilemma opposing complexity and implementation. **It is difficult to update and revamp regulations permanently if one expects these rules to be implemented across the board.** Just as impairing the one-stop shop will weaken the GDPR as a whole, **propositions should focus in priority on clarity, simplicity and ease of implementation.** These will not be enough to resolve fundamental challenges from future technologies – but public support for the GDPR will wane if it is not seen as effective in the short run and on issues which any one can see.

2. Making privacy policies more readable and ergonomic

Some of the needed improvements are obvious. Neither the GDPR nor the ensuing guidelines make much of UX – user experience – in assessing what is a primary goal of the GDPR – putting the individual back in control of his/her personal data.

a. **Standardizing notice and consent pop-up forms.** Any current user soon finds out some are more ergonomic and usable than others. Others are complexified by such intermediary steps as reading privacy policies from various sources. In extreme but frequent cases there is no possibility to decide – instead, this is information of opt-in.

b. **Single- and multi-point data collection requests should be opt-out by default,** failure to respond should signify opt-out, for example. As it is, Recital 32 of the GDPR requires a clear affirmative action for opt-in, but pre-ticked opt-in boxes are not enough. Closing a window or disregarding it should not constitute such a sign.

c. Use of semi-literate icons and codes: it is not only a necessity for a country with high levels of illiteracy. Just as road driving signs are standardized and memorized, **setting up a digital driving code will ease the issue of time-consuming and difficultly-worded privacy policies.**

d. **AI apps facilitating privacy:** some apps make it possible to read and analyze privacy policies. **Reviewing and eventually endorsing or grading these apps, and publicizing the results, would serve the public.** Examples among these apps: *Guard,*[230] a neural app that analyzes the privacy terms of known sites and grades them; Terms of service; Didn't read (ToS; DR),[231] that classifies and rates the fine print of privacy policies from Class A to Class E (from the most to the least protective). It is crowd-sourced and peer-reviewed.

[230] "Discover the Hidden Secrets in Privacy Policies | Guard," *Guard,* https://useguard.com.

[231] "Terms of Service; Didn't Read," *Tosdr,* https://tosdr.org/classification.html.

Beyond these examples, **privacy assistance to users is now a research field that should be publicly funded in the European Union.** An American equivalent is Carnegie Mellon University's *Personalized Privacy Assistant* project,[232] funded by DARPA, the U.S. Air Force, the National Science Foundation, Google and Yahoo. The project is also a member of a wider consortium called *The Usable Privacy Project.*[233]

3. Ensuring privacy by design in practice

a. The GDPR has explicitly recognized data protection by design in Recitals 78 and 108, as well as in Article 25. Privacy by design incorporates data minimization, purpose limitation, retention and incorporation of privacy at a development life cycle's first stage. No formal guidelines have been issued but the European Data Protection Supervisor (EDPS) issued an Opinion in May 2018.[234] It acknowledges some of the research described above, and makes recommendations. Among these: ensuring that **the European Union Agency for Cybersecurity (ENISA) has the resources to encourage more research and interaction with private companies; by "increasing incentives and substantiating obligations, including appropriate liability rules"; "to support an inventory and observatory of the "state of the art" of privacy engineering"; and by "providing guidance to controllers".** These are all excellent

[232] "Personalized Privacy Assistant Project," *Privacyassistant.org* (Carnegie Mellon University, 2018), https://www.privacyassistant.org.

[233] "The Usable Privacy Policy Project," *Usableprivacy.org,* 2019, https://www.usableprivacy.org.

[234] "Preliminary Opinion on Privacy by Design," *The European Data Protection Supervisor,* May 31 2018, https://edps.europa.eu/sites/edp/files/publication/18-05-31_preliminary_opinion_on_privacy_by_design_en_0.pdf.

suggestions, but they are expressed as hopes rather than prescriptions. The reality is that an overwhelming share of the research around privacy by design– whether from the field of social sciences or around digital innovation, originates in the United States –, is often conducted through programs that mix federal support, private firms and scientific establishments. A truly remarkable research by the European Parliamentary Research Service (EPRS) led to a report on algorithmic accountability and transparency. This independent report, whose most salient recommendation is to encourage and protect whistleblowers inside organizations, includes 576 endnotes: they are overwhelmingly sources from American literature.[235]

b. It is therefore necessary **to strengthen the research and the links between rule-makers, business and the scientific community** beyond the level recommended by the Opinion of the EDPS. Companies and compliance advisors unanimously praise the virtues of privacy by design, but they come short on how they actually implement it. One must surmise that there is a tension inside each company between business goals and privacy. By themselves, data protection officers may not be influential enough to ensure a fair balance between these goals. The GDPR emphasizes internal Data Protection Impact Assessments (DPIAs) as a formal process identifying risks, appropriate control and mitigation steps. Their range is narrower than privacy by design, and in this area, there is indeed an earlier Guideline from the Article 29 Working Party that preceded the EDPB.[236] Insurance

[235] "A Governance Framework for Algorithmic Accountability and Transparency - Think Tank" *European Parliamentary Research Service,* April 2019, http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2019)624262.

[236] European Commission, "Guidelines on Data Protection Impact Assessment," 2017, https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236.

companies are perhaps the most familiar with the process, since they handle huge amounts of data that have strong implications for privacy. **The need for clear policies, guidelines and work instructions related to data protection applies to public guidance as well:** whether within the coming e-Privacy directive or through new GDPR guidelines, this is the most pragmatic way to provide companies – and especially their data compliance officers – with a roadmap that goes beyond abstract requirements.

4. Providing an effective right to explanation under the GDPR

a. Recital 71 provides the public a right to explanation for decisions reached by automated process. Algorithms are black boxes to almost all users. For explanation, the largest project so far is Explainable Artificial Intelligence (XAI), an initiative launched in 2016 by the Defense Advanced Research Projects Agency (DARPA),[237] the very agency that led to ARPANET, the precursor of the internet. As one might already expect, the majority of stated goals and programs are not about users' privacy. The project is explicitly aimed at improving the efficiency of deep learning: "machine-learning systems will have the ability to explain their rationale, characterize their strengths and weaknesses, and convey an understanding of how they will behave in the future". In that sense, public explainability to individual users and accountability are only a by-product. Several projects, such as Texas A&M University's on detecting fake news, UC Berkeley's on autonomous vehicles or on acquiring a "reasonably accurate mental model of

[237] "Explainable Artificial Intelligence," *Defense Advanced Research Projects Agency,* https://www.darpa.mil/program/explainable-artificial-intelligence.

robots' policies", and the overall accent on helping users understand the basis for the algorithms they use, would clearly have implications for explanations of decisions to individuals.

Much more mundane examples for explainability can be found with CreditKarma, which provides understanding of credit scores. Other tools have been created such as Google's Match score on Maps[238] or Netflix Percent Match.

b. **Product liability** – warranties, the responsibility to do no harm – were a consequence of the industrial revolution. Inasmuch as digital data is about commerce, and if we admit that the train has left the station where personal data could not be traded, **it is a good model for the digital age as well. Rather than deny the commerciality of much personal data, we should adapt our enforcement process to this reality.**

c. The above-mentioned EPRS report on algorithm accountability has interesting policy suggestions that apply primarily to the public sector and to exceptions based on public interest. The first set applies to **raising awareness through education, watchdogs and whistleblowers.** The report recommends exceptions **allowing reverse engineering of algorithms in cases of public interest.** Important cases, from aspects of the Snowden affair to Propublica's exposition of machine bias[239] are cited – one might add that it is reverse engineering that has allowed for a critical discovery of the

[238] Mariella Moon, "Google Maps Can Predict How Much You'll like a Restaurant," *Engadget*, July 31, 2018, https://www.engadget.com/2018/07/31/google-maps-match-feature/.

[239] Julia Angwin, et al., "Machine Bias," *ProPublica,* May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

software used by public security against Xinjiang's population.[240] But reverse engineering is often prohibited by law or precluded by intellectual property rights, and **whistleblower exemptions should therefore be granted in the public interest.** To this, we should also add a recommendation that **open-source code be used as much as possible for the software programs incorporating algorithms.** Transparency has also become a source of cybersecurity, as opposed to millions of lines of unverifiable coding. The current trend towards open-source coding combines more cybersecurity with more accountability in the public sphere. If there is to be verification of the innocuity of programs, of the absence of malware or snippets inserted to siphon data, open-source algorithms and coding are necessary – if not sufficient by themselves. In France, the Villani report on AI has proposed **the creation of a group of certified public experts.** These could conduct audits of algorithms and databases, and carry out testing; they could be called in during legal proceedings.

Inversely, **for the transmission of private or sensitive personal data, end-to-end encryption is a frequent recommendation,** even if it is not completely tamperproof: it is still vulnerable at both ends, and quantum computing is said to herald the coming day when safe encryption codes come to an end. The counter-argument is of course the need to find out about criminal activity and to prevent it from "going dark". **These debates should not be eluded – but neither should they be taken up only when emergencies occur. Privacy as a goal then recedes behind security.**

[240] "China's Algorithms of Repression | Reverse Engineering a Xinjiang Police Mass Surveillance App," *Human Rights Watch,* May 1, 2019, https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance

In sum, **transparency is required for public algorithms that enable decisions regarding individuals, and open-source software makes it more difficult to hide snooping codes and other malware,** provided of course there is regular supervision. This applies to programs collecting, interpreting and using data. On the contrary, **opacity is required for the transmission and storage of private data, whether it belongs to individuals or to companies. In that case, end-to-end encryption is a desirable approach.** Exceptions due to considerations for the public order should be made only with caution and with the necessary presence of independent oversight.

5. Creating avenues for torts and litigation

The European Union used to assess anti-dumping penalties according to the "lesser duty rule" – the penalties should be only enough to compensate the actual injury caused by dumped goods, rather than be based on the overall amount of dumping.[241] The huge fines imposed in other areas by U.S. regulatory agencies on non-American companies, and the gigantic cash resources of Silicon Valley's IT giants, have inspired a rethink at the European Commission on the size of fines based on violations of competition law. Similarly, the egregiousness of dumping practices by Chinese companies on the EU market have also sparked a debate on the methods for assessing anti-dumping penalties. **European legal and regulatory practices is in a welcome transition from one approach to the other. This is in fact part of a broader trend, but it also applies to the area of**

[241] François Godement, "China's Market Economy Status and the European Interest," *European Council on Foreign Relations,* June 23, 2016, p. 7, https://www.ecfr.eu/page/-/ECFR_180_-_CHINA_MARKET_ECONOMY_STATUS_AND_THE_EUROPEAN_INTEREST_%28002%29.pdf.

privacy violations. In the case of Facebook's 2019 fine, the FTC proudly said this fine was 20 times larger than the next privacy fine worldwide, and that a court would have unlikely to rule in favor of such a change of scale. The size of the fine was meant as a signal to all companies.[242]

The other ex-post action is litigation. The earlier mentioned Hand Formula, which is based on the assessment of proportionality and takes into account the firm's scale, is a good start. The EPRS report has good proposals both on accountability of algorithms and on litigation. The report urges **member states to make efforts for more public accountability on the algorithms they use for decisions.** Examples exist, such as the publication of the tax formula in France. Public reviews and periodic impact assessments, given newly appearing techniques, are called for. The consolidation of personal data from private sources and companies under public contract, and the subsequent number-crunching and algorithmic treatment done by these public agencies on privately collected data, should be a particular subject of attention.

The same requirement of complete transparency **does not hold true for the private sector and companies, in part because it is hard to achieve, and also because it may run against a business' interest: an algorithm is a vital function of a company.** Here, the report moves in an American direction on the key issue of torts and litigation: **"it would be preferable to establish a legal liability framework that allows service providers to accept greater tort**

[242] Lesley Fair, "FTC's $5 Billion Facebook Settlement: Record-Breaking and History-Making," *Federal Trade Commission,* July 24, 2019, https://www.ftc.gov/news-events/blogs/business-blog/2019/07/ftcs-5-billion-facebook-settlement-record-breaking-history.

liability in exchange for reduced transparency and Algorithmic Impact Assessment requirements."[243] Finally, the report extends to global initiatives on what it calls the "Fourth Industrial Revolution (…): AI arms race". It proposes "a strong position in trade negotiations to protect regulatory ability to investigate algorithmic systems and hold parties accountable for violations of European laws and human rights" and builds on the proposition to set up an International Artificial Intelligence Organization on the model of the existing International Telecommunications Organization (ITU). Their inspiration for this: four researchers, of whom three live and work in America and one in the United Kingdom.

We strongly endorse the EPRS proposal to establish stronger ex-post tort while limiting new demands to private companies on algorithmic accountability.

6. Introducing sector-specific regulations

There are clearly sectors that require specific or more fine-tuned regulation. The health sector, financial services, police data are such cases. It is in part achieved with the EU's so-called Police Directive, in fact relating to police and justice, which was adopted in 2016, prior to the GDPR.[244] As a directive, it requires transposition into member state laws, which are open to interpretations. Nonetheless, **cases are currently being proceeded by the CJEU regarding the**

[243] "A Governance Framework for Algorithmic Accountability and Transparency - Think Tank," *European Parliamentary Research Service*, April 2019, p. 73, http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2019)624262.

[244] "Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016," EUR-Lex, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016L0680.

legality of large-scale collection and storage of surveillance data by police and intelligence agencies. They have been introduced by the UK (following a challenge by Privacy International), and by French and Belgian higher administrative courts (following suits from NGOs) as checks on the legality of their national actions. Extensive hearings are underway.[245]

In other sectors, there are the examples of DISHA, the extensive health data regulation bill now under discussion in India, and HIPAA, the 1996 American Health Insurance Accountability and Portability Act. For data protection in the financial sector, the U.S. took the lead as early as 1999 under the wider Gramm-Leach-Bliley Financial Modernization Act. **Encouraging sectoral rules in some cases for the EU** does not mean endorsing the extraordinary maze of state-by-state regulations that it usually ensues.

7. Creating simulated health data for better anonymization

Our in-focus theme, health data protection, reflects a universal concern across systems. In the search for compromises or better solutions, **simulated health data is emerging as a technological way to deal with the failures of anonymization and pseudonymization.** One example has been developed by Simulacrum, backed by major pharma firms, to deal with cancer patient data turned over by Public Health England (PHE). This is especially significant as the UK's

[245] Bill Goodwin, "European Court to Decide on Legality of Bulk Phone and Internet Surveillance," *Bill Goodwin,* September 13, 2019, https://www.computerweekly.com/news/252470666/European-court-to-decide-on-legality-of-bulk-communications-surveillance.

National Health Service has been previously criticized for turning over data to Deepmind, and in effect to Google. With the new technique, data is first anonymized and grouped in batches of minimum 50 samples by PHE. Turned over to Simulacrum, the data is then synthetized in ways that never replicate an actual patient. **This additional layer reconciles the need for big data research with a guarantee for data privacy,** and it deserves to be studied and expanded. Unavoidably, the result is only as good as the algorithm used, and some data details and associations are likely to be lost in the process.

ACKNOWLEDGEMENTS

This exploration of the global digital privacy debate owes much to the generous time granted to the author by Institut Montaigne for research and writing. It would not have been possible without the help of **Meeta Tarani**, intern and program officer, and **Viviana Zhu,** policy officer, who assisted in the research and provided stimulating remarks during writing. At one stage or another, **Gilles Babinet, Eric Chaney, Théodore Christakis, Mathieu Duchâtel, Marie-Anne Frison-Roche, Théophile Lenoir, Angèle Malâtre-Lansac, Laure Millet, Victor Poirier,** and **Stefan Soesanto** have provided useful insights and comments.

People Interviewed

- **Andrea Carrera Marisca**l, Data Protection Legal Counsel, Orange S.A.
- **Mathieu Coulaud**, Head of Legal, Microsoft France
- **Olivier Esper,** Government Affairs & Public Policy Senior Manager, Google France
- **Christophe Fessart,** Data Protection Officer, Enedis
- **Clotilde Jolivet,** Head of Government Affairs France, Sanofi
- **Franck Perraudin,** Head of External Affairs Asia, Sanofi
- **Jean-Renaud Roy,** Director of Corporate Affairs, Microsoft France
- **Ralf Sauer,** Deputy Head of International Data Flows and Protection Unit, DG Justice and Consumers, European Commission
- **Fabien Venries,** Head of Privacy & Marketing Stream, Orange Group

Finally, able proofreading was done by **Paula Martínez López** and **Pierre Pinhas**, and this study has been presented to you thanks to Institut Montaigne's Communication Team.

The opinions expressed in this study are not necessarily those of the above-mentioned persons or the institutions that they represent.

OUR PREVIOUS PUBLICATIONS

- Médicaments innovants : prévenir pour mieux guérir (September 2019)
- Rénovation énergétique : chantier accessible à tous (July 2019)
- Agir pour la parité: performance à la clé (July 2019)
- Pour réussir la transition énergétique (June 2019)
- Europe-Afrique : partenaires particuliers (June 2019)
- Media polarization « à la française »? Comparing the French and American ecosystems (May 2019)
- L'Europe et la 5G : le cas Huawei (Part 2, May 2019)
- L'Europe et la 5G : passons la cinquième ! (Part 1, May 2019)
- Système de santé : soyez consultés ! (April 2019)
- Travailleurs des plateformes : liberté oui, protection aussi (April 2019)
- Action publique : pourquoi faire compliqué quand on peut faire simple (March 2019)
- La France en morceaux : baromètre des Territoires 2019 (February 2019)
- Énergie solaire en Afrique : un avenir rayonnant ? (February 2019)
- IA et emploi en santé : quoi de neuf docteur ? (January 2019)
- Cybermenace : avis de tempête (November 2018)
- Partenariat franco-britannique de défense et de sécurité : améliorer notre coopération (November 2018)
- Sauver le droit d'asile (October 2018)
- Industrie du futur, prêts, partez ! (September 2018)
- La fabrique de l'islamisme (September 2018)
- Protection sociale : une mise à jour vitale (March 2018)
- Innovation en santé : soignons nos talents (March 2018)
- Travail en prison : préparer (vraiment) l'après (February 2018)
- ETI : taille intermédiaire, gros potentiel (January 2018)
- Réforme de la formation professionnelle : allons jusqu'au bout ! (January 2018)
- Espace : l'Europe contre-attaque ? (December 2017)
- Justice : faites entrer le numérique (November 2017)
- Apprentissage : les trois clés d'une véritable transformation (October 2017)
- Prêts pour l'Afrique d'aujourd'hui ? (September 2017)
- Nouveau monde arabe, nouvelle « politique arabe » pour la France (August 2017)
- Enseignement supérieur et numérique : connectez-vous ! (June 2017)
- Syrie : en finir avec une guerre sans fin (June 2017)
- Énergie : priorité au climat ! (June 2017)
- Quelle place pour la voiture demain ? (May 2017)
- Sécurité nationale : quels moyens pour quelles priorités ? (April 2017)

- Tourisme en France : cliquez ici pour rafraîchir (March 2017)
- L'Europe dont nous avons besoin (March 2017)
- Dernière chance pour le paritarisme de gestion (March 2017)
- L'impossible État actionnaire ? (January 2017)
- Un capital emploi formation pour tous (January 2017)
- Économie circulaire, réconcilier croissance et environnement (November 2016)
- Traité transatlantique : pourquoi persévérer (October 2016)
- Un islam français est possible (September 2016)
- Refonder la sécurité nationale (September 2016)
- Bremain ou Brexit : Europe, prépare ton avenir ! (June 2016)
- Réanimer le système de santé - Propositions pour 2017 (June 2016)
- Nucléaire : l'heure des choix (June 2016)
- Un autre droit du travail est possible (May 2016)
- Les primaires pour les Nuls (April 2016)
- Le numérique pour réussir dès l'école primaire (March 2016)
- Retraites : pour une réforme durable (February 2016)
- Décentralisation : sortons de la confusion / Repenser l'action publique dans les territoires (January 2016)
- Terreur dans l'Hexagone (December 2015)
- Climat et entreprises : de la mobilisation à l'action / Sept propositions pour préparer l'après-COP21 (November 2015)
- Discriminations religieuses à l'embauche : une réalité (October 2015)
- Pour en finir avec le chômage (September 2015)
- Sauver le dialogue social (September 2015)
- Politique du logement : faire sauter les verrous (July 2015)
- Faire du bien vieillir un projet de société (June 2015)
- Dépense publique : le temps de l'action (May 2015)
- Apprentissage : un vaccin contre le chômage des jeunes (May 2015)
- Big Data et objets connectés. Faire de la France un champion de la révolution numérique (April 2015)
- Université : pour une nouvelle ambition (April 2015)
- Rallumer la télévision : 10 propositions pour faire rayonner l'audiovisuel français (February 2015)
- Marché du travail : la grande fracture (February 2015)
- Concilier efficacité économique et démocratie : l'exemple mutualiste (December 2014)
- Résidences Seniors : une alternative à développer (December 2014)
- Business schools : rester des champions dans la compétition internationale (November 2014)

- Prévention des maladies psychiatriques : pour en finir avec le retard français (October 2014)
- Temps de travail : mettre fin aux blocages (October 2014)
- Réforme de la formation professionnelle : entre avancées, occasions manquées et pari financier (September 2014)
- Dix ans de politiques de diversité : quel bilan ? (September 2014)
- Et la confiance, bordel ? (August 2014)
- Gaz de schiste : comment avancer (July 2014)
- Pour une véritable politique publique du renseignement (July 2014)
- Rester le leader mondial du tourisme, un enjeu vital pour la France (June 2014)
- 1 151 milliards d'euros de dépenses publiques : quels résultats ? (February 2014)
- Comment renforcer l'Europe politique (January 2014)
- Améliorer l'équité et l'efficacité de l'assurance-chômage (December 2013)
- Santé : faire le pari de l'innovation (December 2013)
- Afrique-France : mettre en œuvre le co-développement. Contribution au XXVI^{e} sommet Afrique-France (December 2013)
- Chômage : inverser la courbe (October 2013)
- Mettre la fiscalité au service de la croissance (September 2013)
- Vive le long terme ! Les entreprises familiales au service de la croissance et de l'emploi (September 2013)
- Habitat : pour une transition énergétique ambitieuse (September 2013)
- Commerce extérieur : refuser le déclin. Propositions pour renforcer notre présence dans les échanges internationaux (July 2013)
- Pour des logements sobres en consommation d'énergie (July 2013)
- 10 propositions pour refonder le patronat (June 2013)
- Accès aux soins : en finir avec la fracture territoriale (May 2013)
- Nouvelle réglementation européenne des agences de notation : quels bénéfices attendre ? (April 2013)
- Remettre la formation professionnelle au service de l'emploi et de la compétitivité (March 2013)
- Faire vivre la promesse laïque (March 2013)
- Pour un « New Deal » numérique (February 2013)
- Intérêt général : que peut l'entreprise ? (January 2013)
- Redonner sens et efficacité à la dépense publique. 15 propositions pour 60 milliards d'économies (December 2012)
- Les juges et l'économie : une défiance française ? (December 2012)
- Restaurer la compétitivité de l'économie française (November 2012)

- Faire de la transition énergétique un levier de compétitivité (November 2012)
- Réformer la mise en examen Un impératif pour renforcer l'État de droit (November 2012)
- Transport de voyageurs : comment réformer un modèle à bout de souffle ? (November 2012)
- Comment concilier régulation financière et croissance : 20 propositions (November 2012)
- Taxe professionnelle et finances locales : premier pas vers une réforme globale ? (September 2012)
- Remettre la notation financière à sa juste place (July 2012)
- Réformer par temps de crise (May 2012)
- Insatisfaction au travail : sortir de l'exception française (April 2012)
- Vademecum 2007 – 2012 : Objectif Croissance (March 2012)
- Financement des entreprises : propositions pour la présidentielle (March 2012)
- Une fiscalité au service de la « social compétitivité » (March 2012)
- La France au miroir de l'Italie (February 2012)
- Pour des réseaux électriques intelligents (February 2012)
- Un CDI pour tous (November 2011)
- Repenser la politique familiale (October 2011)
- Formation professionnelle : pour en finir avec les réformes inabouties (October 2011)
- Banlieue de la République (September 2011)
- De la naissance à la croissance : comment développer nos PME (June 2011)
- Reconstruire le dialogue social (June 2011)
- Adapter la formation des ingénieurs à la mondialisation (February 2011)
- « Vous avez le droit de garder le silence… ». Comment réformer la garde à vue (December 2010)
- Gone for Good? Partis pour de bon ? Les expatriés de l'enseignement supérieur français aux États-Unis (November 2010)
- 15 propositions pour l'emploi des jeunes et des seniors (September 2010)
- Afrique - France. Réinventer le co-développement (June 2010)
- Vaincre l'échec à l'école primaire (April 2010)
- Pour un Eurobond. Une stratégie coordonnée pour sortir de la crise (February 2010)
- Réforme des retraites : vers un big-bang ? (May 2009)
- Mesurer la qualité des soins (February 2009)

- Ouvrir la politique à la diversité (January 2009)
- Engager le citoyen dans la vie associative (November 2008)
- Comment rendre la prison (enfin) utile (September 2008)
- Infrastructures de transport : lesquelles bâtir, comment les choisir ? (July 2008)
- HLM, parc privé. Deux pistes pour que tous aient un toit (June 2008)
- Comment communiquer la réforme (May 2008)
- Après le Japon, la France… Faire du vieillissement un moteur de croissance (December 2007)
- Au nom de l'Islam… Quel dialogue avec les minorités musulmanes en Europe ? (September 2007)
- L'exemple inattendu des Vets. Comment ressusciter un système public de santé (June 2007)
- Vademecum 2007-2012. Moderniser la France (May 2007)
- Après Erasmus, Amicus. Pour un service civique universel européen (April 2007)
- Quelle politique de l'énergie pour l'Union européenne ? (March 2007)
- Sortir de l'immobilité sociale à la française (November 2006)
- Avoir des leaders dans la compétition universitaire mondiale (October 2006)
- Comment sauver la presse quotidienne d'information (August 2006)
- Pourquoi nos PME ne grandissent pas (July 2006)
- Mondialisation : réconcilier la France avec la compétitivité (June 2006)
- TVA, CSG, IR, cotisations… Comment financer la protection sociale (May 2006)
- Pauvreté, exclusion : ce que peut faire l'entreprise (February 2006)
- Ouvrir les grandes écoles à la diversité (January 2006)
- Immobilier de l'État : quoi vendre, pourquoi, comment (December 2005)
- 15 pistes (parmi d'autres…) pour moderniser la sphère publique (November 2005)
- Ambition pour l'agriculture, libertés pour les agriculteurs (July 2005)
- Hôpital : le modèle invisible (June 2005)
- Un Contrôleur général pour les Finances publiques (February 2005)
- Les oubliés de l'égalité des chances (January 2004 - Re-edition September 2005)

For previous publications, see our website:
www.institutmontaigne.org

INSTITUT
MONTAIGNE

ABB FRANCE
ABBVIE
ACCURACY
ACTIVEO
ADIT
AIR FRANCE – KLM
AIR LIQUIDE
AIRBUS GROUP
ALLEN & OVERY
ALLIANZ
ALVAREZ & MARSAL FRANCE
AMAZON WEB SERVICES
ARCHERY STRATEGY CONSULTING
ARCHIMED
ARDIAN
ASTORG
ASTRAZENECA
A.T. KEARNEY
AUGUST DEBOUZY
AVRIL
AXA
BAKER & MCKENZIE
BANK OF AMERICA MERRILL LYNCH
BEARINGPOINT
BESSÉ
BNP PARIBAS
BOLLORÉ
BOUGARTCHEV MOYNE ASSOCIÉS
BOUYGUES
BRUNSWICK
CAISSE DES DÉPÔTS
CAPGEMINI
CAPITAL GROUP
CAREIT
CARREFOUR
CASINO
CHAÎNE THERMALE DU SOLEIL
CHUBB
CIS
CISCO SYSTEMS FRANCE
CMA CGM
CNP ASSURANCES
COHEN AMIR-ASLANI
COMPAGNIE PLASTIC OMNIUM
CONSEIL SUPÉRIEUR DU NOTARIAT
CORREZE & ZAMBEZE
CRÉDIT AGRICOLE
CRÉDIT FONCIER DE FRANCE
D'ANGELIN & CO. LTD
DASSAULT SYSTEMES
DE PARDIEU BROCAS MAFFEI
DENTSU AEGIS NETWORK
DRIVE INNOVATION INSIGHTS - DII
EDF
EDHEC BUSINESS SCHOOL
EDWARDS LIFESCIENCES FRANCE
ELSAN
ELSEVIER SCIENCES
ENEDIS
ENGIE
EQUANCY
ETHIQUE & DEVELOPPEMENT
EURAZEO
EUROGROUP CONSULTING
EUROSTAR
FIVES
FONCIÈRE INEA
GALILEO GLOBAL EDUCATION FRANCE
GETLINK
GIDE LOYRETTE NOUEL
GOOGLE
GRAS SAVOYE
GROUPAMA
GROUPE EDMOND DE ROTHSCHILD
GROUPE M6
GROUPE ORANGE
HAMEUR ET CIE
HENNER
HSBC FRANCE
IBM FRANCE
IFPASS
ING BANK FRANCE
INSEEC
INTERNATIONAL SOS
INTERPARFUMS
IONIS EDUCATION GROUP
ISRP
JEANTET & ASSOCIÉS
KANTAR
KATALYSE

INSTITUT
MONTAIGNE

KPMG S.A.
LA BANQUE POSTALE
LA PARISIENNE ASSURANCES
LAZARD FRÈRES
LINEDATA SERVICES
LIR
LIVANOVA
L'ORÉAL
LOXAM
LVMH - MOËT-HENNESSY - LOUIS VUITTON
M.CHARRAIRE
MACSF
MALAKOFF MÉDÉRIC
MAREMMA
MAZARS
MCKINSEY & COMPANY FRANCE
MÉDIA-PARTICIPATIONS
MEDIOBANCA
MERCER
MERIDIAM
MICHELIN
MICROSOFT FRANCE
MITSUBISHI FRANCE
NATIXIS
NEHS
NESTLÉ
NEXITY
OBEA
ODDO BHF
ONDRA PARTNERS
ONET
OPTIGESTION
ORANO
ORTEC GROUP
PAI PARTNERS
PRICEWATERHOUSECOOPERS
PRUDENTIA CAPITAL
RADIALL
RAISE
RAMSAY GÉNÉRALE DE SANTÉ
RANDSTAD
RATP
RELX GROUP
RENAULT
REXEL
RICOL LASTEYRIE CORPORATE FINANCE
RIVOLIER
ROCHE
ROLAND BERGER
ROTHSCHILD MARTIN MAUREL
SAFRAN
SANOFI
SCHNEIDER ELECTRIC
SERVIER
SGS
SIA PARTNERS
SIACI SAINT HONORÉ
SIEMENS
SIER CONSTRUCTEUR
SNCF
SNCF RÉSEAU
SODEXO
SOFINORD-ARMONIA
SOLVAY
SPRINKLR
STAN
SUEZ
SYSTEMIS
TALAN
TECNET PARTICIPATIONS SARL
TEREGA
THE BOSTON CONSULTING GROUP
TILDER
TOTAL
TRANSDEV
UBER
UBS FRANCE
UIPATH
VEOLIA
VINCI
VIVENDI
VOYAGEURS DU MONDE
WAVESTONE
WENDEL
WILLIS TOWERS WATSON
WORDAPPEAL

Imprimé en France
Dépôt légal : novembre 2019
ISSN : 1771-6756
Achevé d'imprimer en novembre 2019

INSTITUT
MONTAIGNE

BOARD OF DIRECTORS

CHAIRMAN

Henri de Castries

VICE-PRESIDENT

David Azéma Associé, Perella Weinberg Partners
Jean-Dominique Senard Président, Renault
Emmanuelle Barbara Senior Partner, August Debouzy
Marguerite Bérard-Andrieu Directeur du pôle banque de détail en France, BNP Paribas
Jean-Pierre Clamadieu Président du Comité exécutif, Solvay
Olivier Duhamel Président, FNSP (Sciences Po)
Marwan Lahoud Associé, Tikehau Capital
Fleur Pellerin Fondatrice et CEO, Korelya Capital, ancienne ministre
Natalie Rastoin Directrice générale, Ogilvy France
René Ricol Associé fondateur, Ricol Lasteyrie Corporate Finance
Arnaud Vaissié Co-fondateur et Président-directeur général, International SOS
Florence Verzelen Directrice générale adjointe, Dassault Systèmes
Philippe Wahl Président-directeur général, Groupe La Poste

PRESIDENT D'HONNEUR

Claude Bébéar, Fondateur et Président d'honneur, AXA

INSTITUT
MONTAIGNE

THERE IS NO DESIRE MORE NATURAL THAN THE DESIRE FOR KNOWLEDGE

Digital Privacy: How Can We Win the Battle?

"Gentlemen don't read other people's mail." Actually, they sometimes do, legally or surreptitiously. How can privacy be restored to our life?

The protection of privacy has become a pervasive concern. Legally, privacy is expressed in terms of personal data protection, and it is a key focus of data protection regulations, along with data security.

Today, the privacy debate has two matrixes. One is clearly the United States, the mother of all privacy debates and an accepted reference point of our study. The other is Europe with its path-breaking GDPR, our primary focus.

But two other regions are also among the world's largest digital centers: India and China. Their choices regarding digital privacy will influence the competition with our systems and determine how much we can have a unified global data flow. The health sector is our theme-in-focus, where we argue Europeans must be ready for major changes revolutionizing health care, at the expense of their traditional preference for privacy.

This study finally ends with seven specific propositions for improving or revising the GDPR.

Follow us on:

Sign up for our weekly newsletter on: www.institutmontaigne.org

Institut Montaigne
59, rue La Boétie - 75008 Paris
Tél. +33 (0)1 53 89 05 60 - www.institutmontaigne.org

November 2019