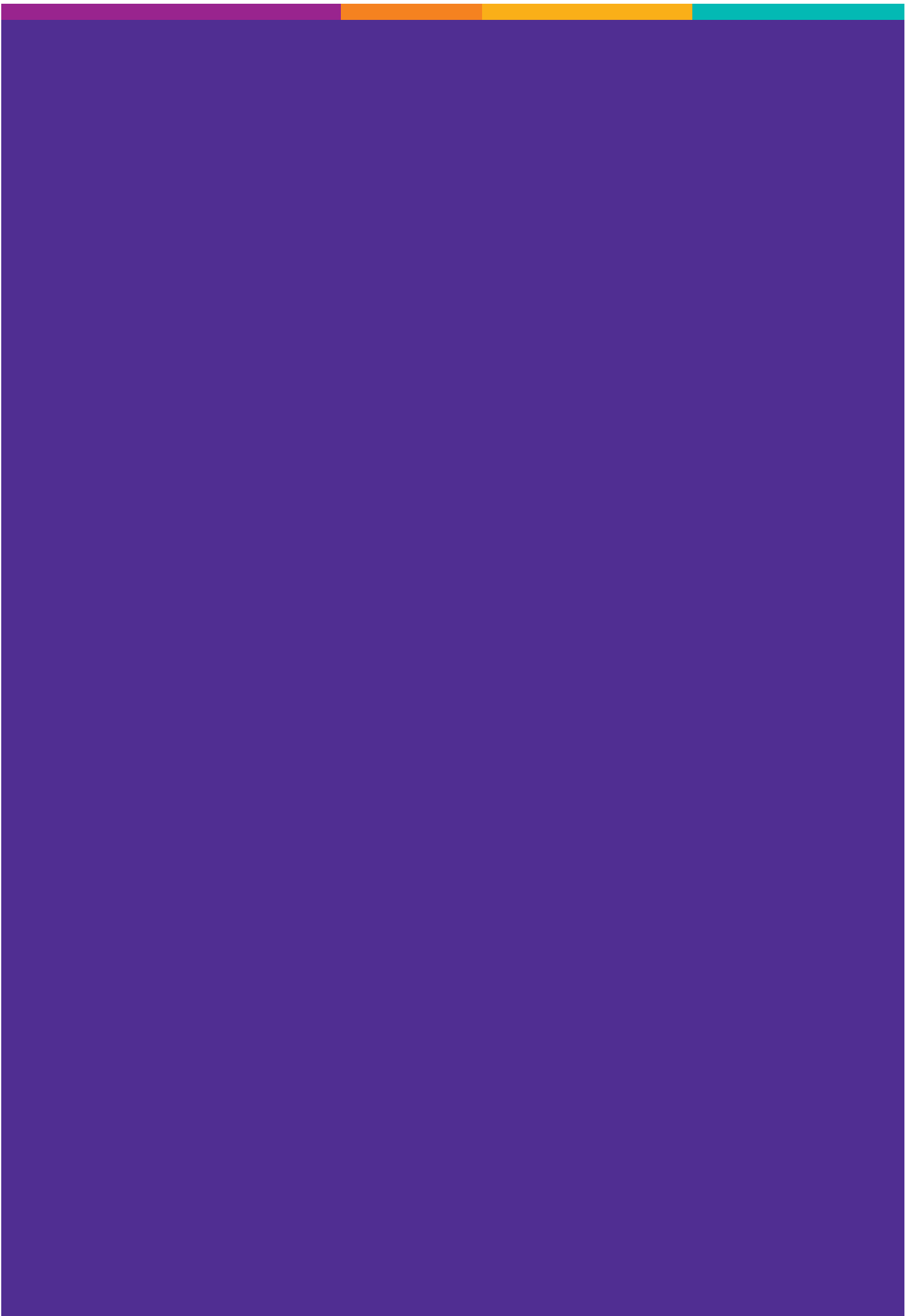


RIO DE JANEIRO 2016

# Big Data in the Global South

*Report on the Brazilian Case Studies*






# Introduction

It is fairly common in the literature about data protection to identify the development of data protection regulation in phases. Viktor Mayer-Schöemberger<sup>1</sup> in fact used to identify generations of data protection laws, the first of them aimed at regulating the few and big computers that were able to manage databases in the late sixties and early seventies, and the fourth of them (and the last documented by this author) aimed at building an efficient enforcement environment with individual informational autodetermination and the assistance of data protection authorities. Data protection regulation, though, faces the crisis in some of its most central and traditional elements - and this crises will most probably move its regulation and the very governance of personal data in a new generation.

Concepts such as the free individual consent or the purpose principle, which were once the central axis of data protection laws, are gradually moving into a complimentary position as more and more personal data is gathered not directly from the individual or even under his knowledge. Big Data, together with other improvements such as those related to the Internet of Things (IoT), are changing considerably the landscape of personal data and will require the regulator to adapt to these new circumstances, whether by adapting existing principles and tools, or by shaping new enforcement means.

Big Data and data protection concerns, of course, are also present in the Global South. In the case of Global South, even if we are dealing with technologies of global reach and also, very often, supplied by global players with similar characteristics all



over the world, some particular issues must be also considered. Countries in the Global South are basically customers and not suppliers of the technology behind Big Data, what can make theoretically these technologies not as fitted to their specific needs.

Moreover, Big Data and its chain of consequences - social sorting, algorithmic decision-making processes and so on - can be used to produce discrimination or even to make some social and economic barriers in some countries' societies even stronger. And it is more like to happen in countries which do not have adequate regulation on subjects related to data protection and digital rights, as is the case of many countries in the region, or when the political landscape is one that can profit from technologies to build systems of surveillance and control for political reasons. So, even if we are dealing with a technological paradigm (Big Data) that we could consider as global, there are reasons for assume that its implementation in Global South must be analysed through some specific criteria.

It is also relevant that Edward Snowden's revelations positioned Brazil, the largest economy in South America, in the middle of the Snowden scandal, after some documents showed that National Security Agency (NSA) was eavesdropping Brazilian's president phone calls.<sup>2</sup> It was not by chance that Brazil, together with Germany – another country whose chancellor was the victim of eavesdropping – presented a proposal of a resolution on privacy online to the United Nations General Assembly, which was unanimously approved<sup>3</sup> and led to the delivering by the UN High Commissioner for Human Rights of a report regarding the right to privacy in the digital era and to the creation of a mandate for a Special Rapporteur on the Right to Privacy,<sup>4</sup> which was recently appointed.<sup>5</sup>

Despite the use of Big Data in Brazil is quite recent, the amount of the data processed in Brazil (and in the Southern Hemisphere) is as impressive as the one of the Northern Hemisphere, but Brazil – different from other South Countries, even Latin American ones – has no general law on data protection, as will be seen in this report. Such a huge amount of data being processed is a consequence also of the use, by local users, of web tools and platforms globally available, such as Google, Youtube and Facebook. Just to give an example, Facebook stores 111 MB of users information on average,<sup>6</sup> and, according to a former Google's President, Eric Schmidt, 5 exabytes are created everyday on the internet, an amount that corresponds to the whole amount of information generated by our civilization since its beginning until 2003.<sup>7</sup>

Besides the collection of data through the use of web platforms and tools, data are also being collected and stored through a wide variety of technological tools. To give an example, just in the City of Rio de Janeiro "it is estimated that there are about 700,000 cameras installed in the streets, buildings, condominiums, banks, supermarkets, etc, that somehow record our daily lives."<sup>8</sup> As highlighted by some authors "we are increasingly leaving traces of our everyday lives."<sup>9</sup>



In this scenario a series of initiatives using big data have also flourished in the south hemisphere. Vast numbers of datasets are being published online by governments pursuing open government principles; Startup and well-established companies are looking for new data to promote new apps, or to manipulate information to sponsor profitable products; even citizens are generating big data by increasingly using more apps, websites, and software to keep their records in the cloud. Big data represents certainly several opportunities, but this trend is not free of concerns. Amongst the areas of concern relating to big data and eventual opportunities and harms, we can name law enforcement, security, public health, transportation, consumer rights and others. Big data promises benefits to a more open society, but also poses risks, both already familiar (such as protection of privacy) and new. An issue of concern in the use of big data is the fact “that access to so much data, from so many different sources, and to the computing power necessary to process it, increasingly means we can perceive patterns, engage in discoveries, and discover secrets that were heretofore hidden”.<sup>10</sup>

Considering this scenario, the present report will present two case studies conducted by the Institute for Technology & Society (Instituto de Tecnologia e Sociedade) on the use of big data for law enforcement purposes in Brazil, one regarding the Federal Police and the other related to the Tax Revenue Service and the Central Bank. The analysis will start by defining the scope of what should be considered as Big Data, then it will focus on the concerns arising from the use of Big Data.

Afterwards, the analysis of the two case studies will be presented, highlighting the best practices, benefits and potential impacts of each case study. The analysis also considered the possible impact of the adoption of a general data protection law in Brazil.

---

<sup>2</sup> See <http://www.globalpost.com/dispatch/news/regions/americas/brazil/130709/us-spying-brazil-snowden-leaks>

<sup>3</sup> UN GA Resolution 68/167. The right to privacy in the digital age. Available at [http://www.un.org/ga/search/view\\_doc.asp?symbol=A/RES/68/167](http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/68/167). Accessed 25 November 2015.

<sup>4</sup> See <http://www.ohchr.org/EN/Issues/Privacy/SR/Pages/SRPrivacyIndex.aspx>. Accessed 25 November 2015.

<sup>5</sup> See <http://www.ohchr.org/EN/Issues/Privacy/SR/Pages/JoeCannataci.aspx>. Accessed 15 December 2015.

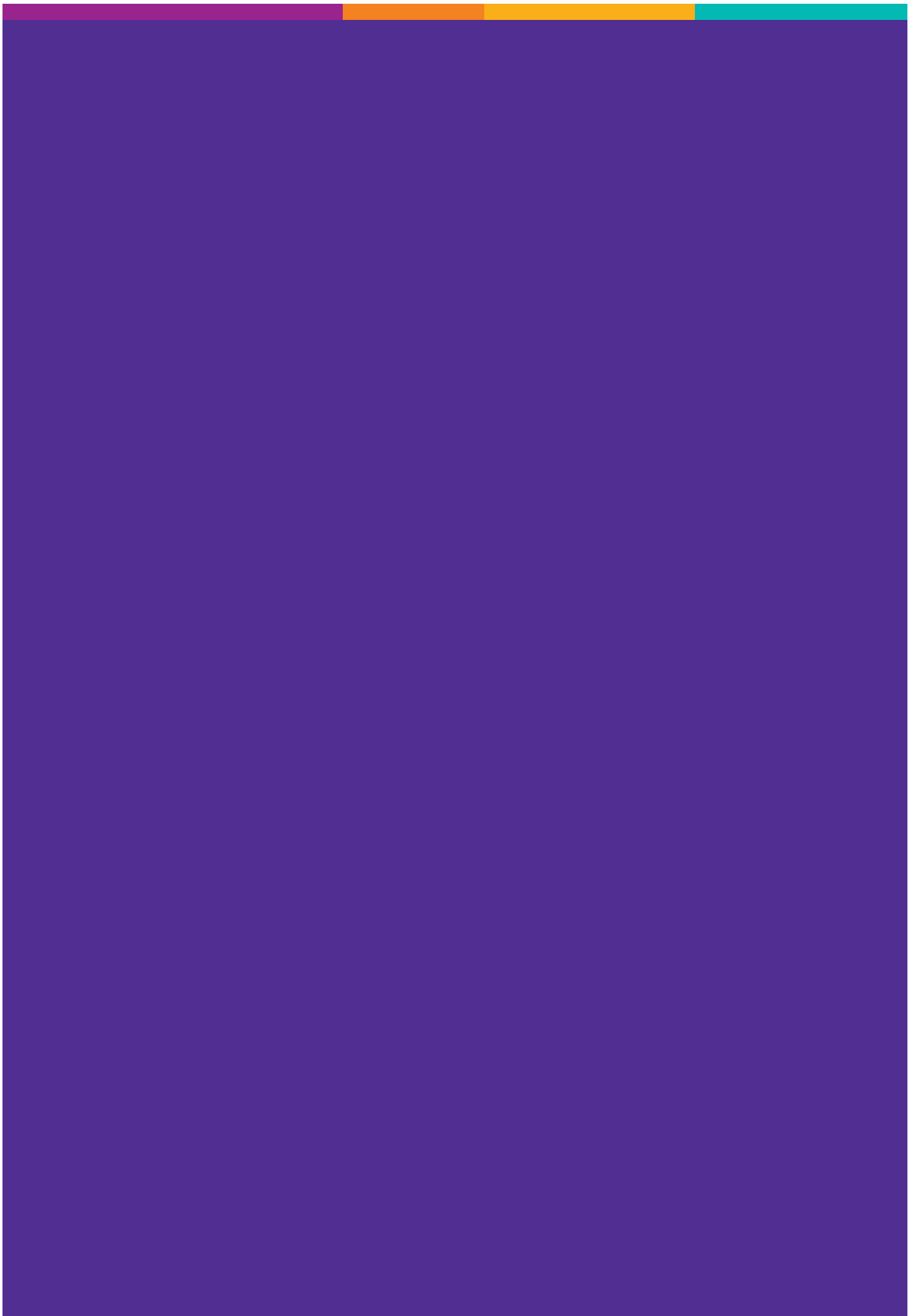
<sup>6</sup> See [https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/privacidade\\_em\\_tempos\\_de\\_big\\_data?lang=en](https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/privacidade_em_tempos_de_big_data?lang=en)

<sup>7</sup> See <http://techcrunch.com/2010/08/04/schmidt-data/>. Accessed 25 November 2015.

<sup>8</sup> See [https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/privacidade\\_em\\_tempos\\_de\\_big\\_data?lang=en](https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/privacidade_em_tempos_de_big_data?lang=en)

<sup>9</sup> See, for instance, LEMOS, André. *Cibercultura e Mobilidade: a Era da Conexão*. Razón y Palabra, N. 41, Oct/Nov 2004. Available at <http://www.razonypalabra.org.mx/antiores/n41/alemos.html>

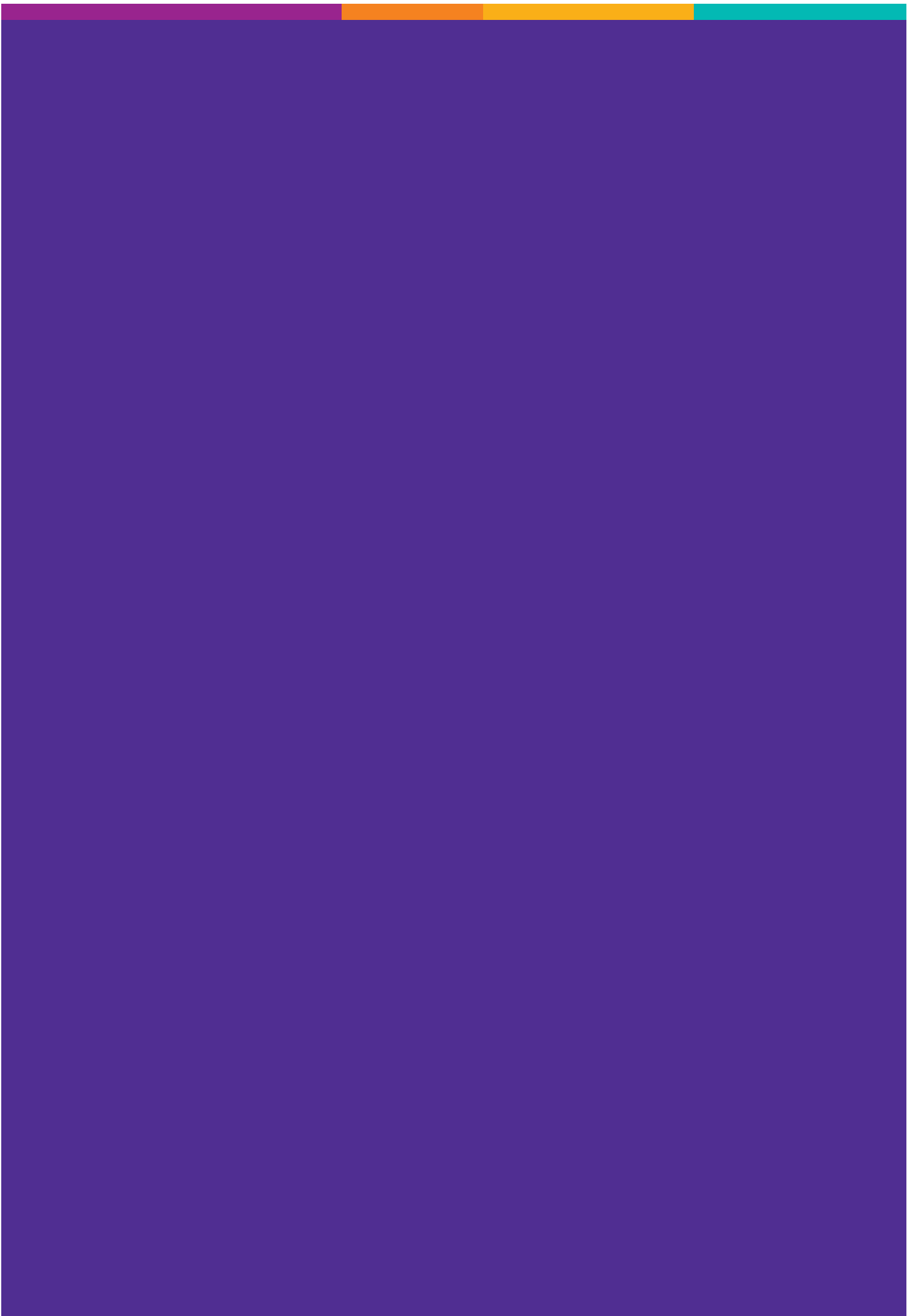
<sup>10</sup> KUNER, Christopher et al. The challenge of ‘big data’ for data protection. Editorial. *International Data Privacy Law*, 2012, Vol. 2, No. 2.





# Summary

<b>Introduction</b>	3
<b>1. What is Big Data</b>	9
<b>2. Concerns</b>	10
<b>3. Passenger Name Record (PNR) and Advance Passenger Information (API) systems</b>	11
<b>3.1. Big data practices within PNR and API</b>	13
<b>3.2. Benefits</b>	14
<b>3.3. Potential implications</b>	15
<b>4. Tax Revenue Service and Central Bank systems: CCS, SCR, and cross-referencing</b>	17
<b>4.1. Big data practices within PNR and API</b>	19
<b>4.2. Benefits</b>	20
<b>4.3. Potential implications</b>	21
<b>5. Possible impact of forthcoming regulation</b>	22
<b>Conclusion</b>	24







# 1/

# What is Big Data?

One of the most often-repeated buzzwords of the last five years is “big data”. This concept, a somewhat natural product of the increasingly rapid development of technology and business models dependent on technology, has proven to be an invaluable asset in a large array of fields, from scholarly research to business analytics and even public policy. But what is big data?

Put simply, it can be said that big data is literally those sets of data, whose existence is possible solely as a consequence of the massive data collection that has become widespread in recent years, thanks to the ubiquitous presence of devices and sensors in everyday life, and the increasing number of people connected to such technologies through digital networks as well of sensors. All actions and communications in digital platforms, such as with mobile phones, computers, or even credit card transactions and, more recently, income tax declarations, or actions that are at some point digitized and thus transformed in data, such as CCTV cameras coupled with facial or pattern recognition software<sup>11</sup>, are prone to be

stored, processed, copied and distributed almost instantaneously, allowing for data analyses that may lead to presumably more well-informed decision making by governments and businesses alike.

While part of this data is collected without the express consent of the data subjects about whom information is generated (such as passenger’s records in international flights), much of it is made available and provided by data subjects’ themselves - through the use of social media services, online shopping, and basically anything one does online that may even be connected to their identity<sup>12</sup>. However, such large-scale data mining and analyses, despite the possible benefits that may result from them, raise many concerns, and those related to privacy and personal data protection top the list. As has been shown time and again, in the era of “big data”, anything you say and do can and probably will be used against - or to help - you<sup>13</sup>.

---

<sup>11</sup> See “UK, the world’s most surveilled state, begins using automated face recognition to catch criminals” <http://www.extremetech.com/extreme/186435-uk-the-worlds-most-surveilled-state-begins-using-automated-face-recognition-to-catch-criminals>

<sup>12</sup> See China’s plan to create a mandatory citizen score system based on social media and shopping behaviour <https://www.aclu.org/blog/free-future/chinas-nightmarish-citizen-scores-are-warning-americans>.

# 2/ Concerns

It has been argued that as more information about individuals is collected by different actors and in different contexts, individual autonomy is undermined: one often does not consent expressly to its data being collected (e.g. in the context of internet browsing analytics) and even when one is given the option of either making use of a service and having their data collected or giving up its use altogether, this choice is made through the use of virtually unreadable terms and conditions and privacy policies.<sup>14</sup> According to Tene and Polonetsky (2012)<sup>15</sup>:

*The harvesting of large data sets and the use of analytics clearly implicate privacy concerns. The tasks of ensuring data security and protecting privacy become harder as information is multiplied and shared even more widely around the world. Information regarding individuals' health, location, electricity use, and online activity is exposed to scrutiny, raising concerns about profiling, discrimination, exclusion, and loss of control.*

Not only is information being collected with weak or completely nonexistent notion of consent, the handling of such information is frequently done with no public transparency or accountability whatsoever, and shared, sold and transmitted to third-parties. Even in the case of purportedly anonymized datasets, it has been shown<sup>16</sup> that re-identification may defeat many attempts at de-identification, hinting that even “anonymous” big data have privacy-related issues.

Privacy concerns are especially alarming when dealing with government data collection and mass

surveillance: as the uncountable scandals involving the NSA and the GCHQ have shown, when left unchecked, State power may be equally or even more harmful to individual liberty and privacy than market initiatives. Though the use of big data by the State may undoubtedly lead to greater efficiency and even more safety, it is vital that both the steps related to data itself - its collection, handling, analysis, etc. - and the decisions made based on such data be subject to public scrutiny and transparency, and be executed with utmost care for the privacy of those with whose data the government deals. That is however often not the case, as will be shown with two case studies on initiatives undertaken by the Brazilian government. In both analyses, we will outline the system's implementation, how its practices are related to the concept of big data, and finally its possible benefits and implications on privacy and data protection.

Focusing on this set of concerns, two situations of the use of Big Data by the Brazilian government were chosen as case studies. Both are solutions that are not strongly documented and that tends to leave concerns about data protection and transparency behind in spite of the technocratic argument of efficiency and the interests of the public administration. In their description some of their potential harm will be unfold.

# 3/ Passenger Name Record (PNR) and Advance Passenger Information (API) systems

The first of such systems consists in fact of two sister systems: the Passenger Name Record (PNR) and the Advance Passenger Information (API). Introduced by the Brazilian National Civil Aviation Agency (ANAC) in 2012 through its Resolution 255<sup>17</sup>, these systems work in parallel by obligating airlines<sup>18</sup> to store and transmit a wide range of data on each and every international flight, its passengers and crew members coming in or out of or simply with a stopover in Brazilian territory.

The information collected must be electronically transmitted to the Federal Police Department (DPF) before each flight - so, passenger's personal data is collected before the flight arrives in Brazil. Officially, the system aims to "prevent and suppress illegal actions", including, for example, tax evasion on imported goods or wanted felons, as well as facilitate entry processes at multiple bureaucratic levels. There are also plans to

<sup>13</sup>An example of this potential use of publicly available personal information is the intention of the US Department of Homeland Security of incorporating social media posts into visa reviews. See <http://www.theverge.com/2015/12/14/10124498/homeland-security-social-media-visa-review>.

<sup>14</sup>See Alexis Madrigal's article to the The Atlantic: "Reading the Privacy Policies You Encounter in a Year Would Take 76 Work Days" <http://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/>

<sup>15</sup>See Tene, O. and Polonetsky, J., "Privacy in the age of big data: a time for big decisions" [http://www.stanfordlawreview.org/sites/default/files/online/topics/64-SLRO-63\\_1.pdf](http://www.stanfordlawreview.org/sites/default/files/online/topics/64-SLRO-63_1.pdf)

<sup>16</sup>See Felten, E. W., Huey, J. and Narayanan, A. "A Precautionary Approach to Big Data Privacy" <http://randomwalker.info/publications/precautionary.pdf>. Accessed 15 December 2015.

broaden the security framework to which these systems belong, for example with the adoption of biometric facial recognition technology in airports, capable of matching passengers' faces with a database of "high risk" individuals.<sup>19</sup>

Moreover, PNR systems have also been implemented or proposed in both the US<sup>20</sup> and the EU<sup>21</sup> and over 50 other countries, and are recommended by the International Civil Aviation Organization (IATA).<sup>22</sup>

---

<sup>17</sup> See ANAC Resolution 255/12:

[http://www2.anac.gov.br/transparencia/audiencia/aud22\\_2012/3%20-%20Resolucao%20-%20API%20e%20PNR%20\(versao%20final\).pdf](http://www2.anac.gov.br/transparencia/audiencia/aud22_2012/3%20-%20Resolucao%20-%20API%20e%20PNR%20(versao%20final).pdf)

<sup>18</sup> Except for services like private jets or helicopters - in Portuguese, "air taxi" (táxi aéreo) services.

<sup>19</sup> See "Receita Federal lança declaração eletrônica de bens de viajantes":

[http://www.receita.fazenda.gov.br/AutomaticoSRFsinot/2013/08/16/2013\\_08\\_16\\_13\\_09\\_50\\_734484890.html](http://www.receita.fazenda.gov.br/AutomaticoSRFsinot/2013/08/16/2013_08_16_13_09_50_734484890.html)

<sup>20</sup> See <http://www.cbp.gov/document/forms/passenger-name-record-pnr-privacy-policy>

<sup>21</sup> See <https://euobserver.com/justice/130430>

<sup>22</sup> See "ANAC determina regras sobre repasse de dados de passageiros à Polícia Federal" <http://economia.uol.com.br/ultimas-noticias/fomoney/2012/11/19/anac-determina-regras-sobre-repasse-de-dados-de-passageiros-a-policia-federal.jhtm>

# 3.1/ Big data practices within PNR and API

In 2013 alone, 19.2 millions of passengers went through Brazil in international flights, a number that is apparently still increasing.<sup>23</sup> The API and PNR systems purport to collect and store detailed, sensitive and private information from all such passengers (as well as their luggage mass and number of bags).

More specifically, over fifteen fields of information per passenger, including full name, gender, home and billing address, birthdate and place and even credit card information. That alone constitutes an impressive amount of information, and the fact that such information is actively shared with other agencies, thus allowing for cross-referencing with other specific databases, is a typical example of big data usage by the public sector.

---

<sup>23</sup> See “ANAC divulga Anuário do Transporte Aéreo de 2013”: [http://www.anac.gov.br/Noticia.aspx?ttCD\\_CHAVE=1584](http://www.anac.gov.br/Noticia.aspx?ttCD_CHAVE=1584)

## 3.2/ Benefits

PNR and API systems aim to make the jobs of government agencies easier. By collecting personally identifiable information on every single passenger coming through Brazil and unifying access to currently sparse databases, government agencies are able to automatically seek passengers in lists of wanted felons, persons of interest, help detect tax evaders, and accelerate the entry process in the country - that is especially true given the implementation of an electronic system for declaration of imported goods for travelers (so called e-DBV)<sup>24</sup>.

And without a doubt, when one does not consider the privacy or data protection implications, it is easy to see how more data about passengers can help investigations, both criminal or not.

---

<sup>24</sup> See "Receita Federal lança declaração eletrônica de bens de viajantes"

[http://www.receita.fazenda.gov.br/AutomaticoSRFsinot/2013/08/16/2013\\_08\\_16\\_13\\_09\\_50\\_734484890.html](http://www.receita.fazenda.gov.br/AutomaticoSRFsinot/2013/08/16/2013_08_16_13_09_50_734484890.html)

# 3.3/ Potential implications

Nearly all types of information that are classified as Personally Identifiable Information under the definition used by the US Dept. of Commerce's National Institute of Standards and Technology are present in PNR records. Therefore, both the amount and nature of the data collected alone are sensitive enough to call for a more robust protection framework. However, although the public tender<sup>25</sup> the Brazilian Federal Police opened to hire contractors for implementing PNR briefly states that such systems must conform to its security guidelines<sup>26</sup>, these are generic at best: no information could be found whatsoever regarding crucial points, and furthermore several other privacy concerns arise upon analysis of the sparse legal framework supporting it. Important points are outlined below, and often coincide with those listed by the IATA as important when implementing PNR systems<sup>27</sup>:

**Purpose limitation:** the nature of the data collected is very personal in nature. From passenger gender to their birth date and credit card number, private and sensitive information is collected with no demonstrated need. Though public safety is important, the mandatory accumulation of personal data of millions of individuals with no control or criteria seems like a blatant attack on personal privacy: a prime example of disproportionality. Even the effectiveness of such a measure is questionable when one considers its selectiveness: as was shown, the transmission of such data is compulsory only for the public at large - and not for private aircrafts.

**Data access clearance:** though it is stated that PNR and API data must be "securely" transmitted and stored, it is not clear who, that is, which agencies and which personnel, will actually have clearance to access such data. It is unclear if connections with other databases will be made automatically by some sort of algorithm, or by hand, and even which databases are in fact connected to such system. Remarks are made that the Brazilian SFR has access to PNR records, and also that they may be used to facilitate clearance with public health and substance control authorities (such as ANVISA), but exactly how such access is made and controlled is unknown.

**Length of information storage period:** the growth rate of the size of such data is enormous. Its storage for an indefinite period of time does not appear to be justified, even if one considers the alleged goals of safekeeping and tax evasion detection. However, at no point there is a reference to the amount of time (if at all) during which that PNR and API data would have to be kept was found.

**Disclosure of individual data:** individuals whose information is collected should be able to obtain a copy of such information, both for transparency as well as for them to be able to make amends to such data. Even though this constitutes, broadly, a constitutional right under Brazilian Law<sup>28</sup> and even can be obtained by means of a FOI (Freedom Of Information) request<sup>29</sup>, at no point in the legislation such a right is described.

Without access to one's data, one is left with no possibility of addressing potential inaccuracies in collected data that may have bad consequences to oneself.

---

<sup>25</sup> See <http://www.dpf.gov.br/servicos/licitacoes/2013/distrito-federal/orgaos-centrais/cgti/pregoes/pregao-eletronico-no-07-2013-cgti-dpf>

<sup>26</sup> See Federal Police Department's "Portaria. No.779/2009-DG/DPF"

<sup>27</sup> See IATA "Guidelines on Passenger Name Record (PNR) Data" [https://www.iata.org/iata/passenger-data-toolkit/assets/doc\\_library/04-pnr/New%20Doc%209944%201st%20Edition%20PNR.pdf](https://www.iata.org/iata/passenger-data-toolkit/assets/doc_library/04-pnr/New%20Doc%209944%201st%20Edition%20PNR.pdf)

<sup>28</sup> Citizens have the right to file a habeas data before Brazilian courts, a kind of procedural document through which one can demand access to government-held information about oneself.

<sup>29</sup> Brazil's legislation on Access to Information (Law 12.527 of 2011) can be used by individuals to request their own personal information, according to its procedure for requests for information.



4/

# Tax Revenue Service and Central Bank systems: CCS, SCR, and cross-referencing

The second collection of systems analyzed is composed of several databases and cross-referencing set-ups utilized by the Brazilian Secretary of Federal Income (SRF, “Secretaria da Receita Federal”) and the Brazilian Central Bank (BCB, “Banco Central do Brasil”) for the purposes of tax control. Through a complex regulatory framework, tax authorities have an enormous range of sources to draw upon during their investigations, that can be divided roughly into two categories:

## A. Financial system tracking:

This includes several databases maintained by the Central Bank and the SFR that collect information on banking operations. Among these are:

I. National Financial System Client Registry (CCS): Introduced by 2003 by law n. 10.701, that enacted changes on previously existing money

laundering laws, the CCS enables the Central Bank to keep records on every single bank account in the country, comprising data on i. client identification, ii. institutions where they maintain accounts or assets and iii. start and termination dates for each business relationship. This does not include sensitive banking information detailing transactions or account movements.

II. Central Bank’s Credit Information System (SRC): a centralized database detailing every credit loan operation exceeding 1000 Brazilian reais (as of Dec/2015, around 264 US\$). The system stores both clients in default or not, and grants each of them a credit score based on how well they pay.

III. Third-party notices or declarations are also a common method: for example, through Normative Instruction RFB N° 811<sup>30</sup>, the Brazilian SFR mandates that banks send them detailed

monthly reports on financial transactions of all accounts that exceed 5000 (for private individuals) or 10.000 Brazilian reais (for legal entities) in a six-month period. This declaration is called the Financial Transaction Information Declaration (Dimof), and it is normatively based on Complementary Law n. 105/2001, regulated by Presidential Decree 4489/2002<sup>31</sup>. There are other similar declarations as well, issued by credit card companies (DECRED), real estate agencies (DIMOB), and even health sector institutions (DMED).<sup>32</sup>

### **B. Outside databases cross-referencing:**

The government also makes use of database integration as well as social network monitoring during its tax audit activities:

**I. Social network analysis:** Brazilian SFR secretary Jorge Rachid has publicly declared this year<sup>33</sup>

that the agency routinely monitors taxpayers' social network profiles in search of incongruences between their tax declarations and their publicly available information. As of now, it is unknown if such practice is institutionalized or is done sporadically and no regulatory text has been found on it.

**II. Access to assorted property record databases:** the SFR also has direct or indirect access to several other databases and sources of information. Particularly, access is granted to national databases of motor vehicles (RENAVAM),<sup>34</sup> aircraft (RAB)<sup>35</sup> and naval vessels (joint with port authorities).<sup>36</sup> Internationally, there is a cooperation agreement with US authorities for information exchange as well.<sup>37</sup>

---

<sup>30</sup> See <http://normas.receita.fazenda.gov.br/sijut2consulta/link.action?visao=anotado&idAto=15765>

<sup>31</sup> See [http://www.planalto.gov.br/ccivil\\_03/decreto/2002/D4489.htm](http://www.planalto.gov.br/ccivil_03/decreto/2002/D4489.htm)

<sup>32</sup> See <http://exame.abril.com.br/seu-dinheiro/noticias/os-dedos-duros-que-entregam-quem-burla-o-imposto-de-renda>

<sup>33</sup> See <http://www.correiobraziliense.com.br/app/noticia/ir2015/2015/05/01/interna2-ir0215,481607/receita-federal-monitora-redes-sociais-dos-contribuintes.shtml>

<sup>34</sup> See <http://normas.receita.fazenda.gov.br/sijut2consulta/link.action?visao=anotado&idAto=20559>

<sup>35</sup> See [http://www.anac.gov.br/Noticia.aspx?ttCD\\_CHAVE=137](http://www.anac.gov.br/Noticia.aspx?ttCD_CHAVE=137)

<sup>36</sup> See <http://exame.abril.com.br/seu-dinheiro/noticias/os-dedos-duros-que-entregam-quem-burla-o-imposto-de-renda>

<sup>37</sup> See <http://idg.receita.fazenda.gov.br/noticias/ascom/2015/agosto/acordo-brasil-eua-permitira-troca-de-informacoes-sobre-contribuintes>

# 4.1/ Big Data practices

While the amount and variety of sources utilized by the SFR certainly approximates their techniques and common big data approaches, it is not clear whether all of such systems or databases are computationally interconnected or if they are separately checked manually, according to atypical demands.

Nevertheless, the fact that the SFR has intensified its efforts both in accumulation of (through an increase in the number of different tax declarations) and access to data, coupled with its heavy investments in technology<sup>38</sup> signal at the very least a trend towards the adoption of big data practices in tax audit.

---

<sup>38</sup> Recently, for instance, 15mi Brazilian reais were spent on facial recognition software in airports. See <http://agenciabrasil.ebc.com.br/economia/noticia/2014-09/receita-reforcara-fiscalizacao-de-passageiros-de-voos-internacionais-em>

## 4.2/ **Benefits**

Undoubtedly, the access to more data generally, and more reliable data especially, make for a much easier and efficient job for both the tax audit and crime fighting duties of the Brazilian Central Bank and SFR. The ability of automatically (that is, computationally) detecting tax evaders, instead of relying on manual labor of comparison and analysis of paperwork makes tax audit faster and swifter.

# 4.3/ Potential implications

Efficiency, however, is not an end by itself, and cannot erase the importance of the rights to privacy and banking secrecy, especially when, as is the case, such practices are done with little to no transparency or accountability. The resource to Big Data must be also followed by considerations about its legal and also ethical implications, otherwise it risks affect the balance of power between the citizen and the public administration in such a way that can effect even the democratic nature of a society. Thus, some points about the potential implications of the use of big data in the Tax Revenue Service and Central Bank systems are outlined below:

**A. Transparency:** Though the myriad of existing tax declaration types are well regulated - by nature, since they impose clear duties of information on third-parties -, many of the other practices that involve database access are poorly, if at all, documented, and are only marginally described in media outlets or press statements. Apart from the SFR access to vehicle property registries, no mention of official access to aircraft and vessel ownership databases can be found in the regulatory *corpus*<sup>39</sup> of the Brazilian SFR. One cannot know, for example, the full extent of personal data being re-transmitted from such databases to the SFR. More worrisome, however, is the allegation of social network surveillance by tax auditors. There is absolutely zero information regarding this subject in any official source, suggesting that such practice is done unofficially, *ad-hoc*, and against the legality

and publicity principles (see article 37 of the Brazilian Republic Constitution).

**B. Limits of the resource to data:** It is unknown the full extent of the surveillance being done, e.g. whether it makes use only of publicly available information, whether such information is stored and used as evidence later or simply as a hint that leads to an official, documented investigation, etc. Still on the matter of tax declarations, there is also debate on the constitutionality of several measures adopted by the SFR in particular and the federal government in general, specifically regarding Complementary Law n. 105/2001. This law grants the Executive branch discretionary power to access banking and financial records with no judicial oversight and even if this can theoretically be at least debatable, there is no jurisprudential formal position on this issue.

**C. Access:** Even if Brazilian Central Bank began to design systems able to give citizens' access to parts of their own personal financial information (for example the Registrato<sup>40</sup> system), tools and provisions to individual's access to personal information lacks in every new conceived system. What looks like is that individual's access is not the rule and should be implemented only if there is a particular justification. Also, individuals who claim access must follow procedures which are not specifically designed for accessing data, what can be slow and bothersome.

<sup>39</sup> Available at Normas SRF: <http://normas.receita.fazenda.gov.br/sijut2consulta/consulta.action>

<sup>40</sup> <http://www.bcb.gov.br/pt-br/sfn/registrato/Paginas/default.aspx>

# 5/ Possible impact of forthcoming regulation

Considering current Brazilian legislation on data protection, there are currently no directly applicable legislation that could serve as a basis for such a treatment of data. In the light of the proposed Draft Bill on data protection, presented by Brazilian's Ministry of Justice in October 2015, it is arguable that PNR and API system could be classified<sup>41</sup> in the category of Art 4°, III:

*Art. 4 - The present law does not apply to data processing:*

...

*III - undertaken for the exclusive purposes of public security, national defense, State security, investigation and law enforcement activities.*

The subject is, therefore, out of the reach of the proposed data protection legislation, except for a consideration that follows in the same article 4°, § 1st of the draft:'

*§ 1st - The data processing provided in item III above will be regulated by a specific law, pursuant to general principles of data protection and to the data subjects' rights as provided by the present Law.*

Such 'principles' and 'data subjects' rights' are further described in the proposed regulations.

Those includes rights such as the access to information to be exercised by citizens (Art. 8° of the proposed bill), as well as other dispositions that should enforce the citizens' position and don't conflict with the purpose of the systems.

As noted before, individual access to one's own data is granted by the Habeas Data writ as well as its procedure is implemented by means of the Access to Information Law (Law 12.527 of 2011); however, these aren't procedures tailored for contemporary data protection problems (for instance, the former can only be requested by a lawyer and the second can only be used against public offices).

Another possible influence of the proposed regulation is the definition it gives to terms such as 'personal data', 'sensitive data' and 'anonymized data', in its article 5°, all of them present in the systems overviewed. Other possible impact would exist in the case a DPA is created by the proposed bill. In this situation, the DPA could, according to the same article 4° of the proposed legislation, issue recommendations or ask controllers for a Privacy Impact Assessment:

*§ 3rd - The competent public body will issue technical*

*opinions and recommendations relating to the exceptions provided for in items II and III, as well as may request to the data controller privacy impact reports.*

---

<sup>41</sup> Available at Normas SRF: <http://normas.receita.fazenda.gov.br/sijut2consulta/consulta.action>

# Conclusion:

While the collection and treatment of personal data by means of the systems examined can roughly be linked to a relevant public interest – and this is indeed the justification for their implementation, fact is that they raise several data protection issues.

First and foremost, there are some points in its implementations where the lack of transparency is concrete – for example, in the absence of provisions regarding citizens' access to their own data or, even more worryingly, in the context of the resource to social networks by Brazilian SFR, an activity that was only mentioned with no hint about the methods and criteria applied. Moreover, this practice hints that there may be other sources of personal data which may be of interest of SRF that are or, in the future, will be used, which calls for increased transparency.

Another point is the consideration of the purpose and necessity of the data collected related with the aim of the treatment. Even if the consideration of a relevant public interest has to be taken into account, there may be issues regarding the secondary use of data (for example, of data that may be considered confidential) or even of data that may be deemed not necessary for the aimed purposes.

However, the lack of a general data protection legislation in Brazil makes considerations such as these, regarding transparency, purpose and necessity, not so obvious as those principles are not yet entrenched into the country's legal framework. So, it becomes clear that the enactment of a data protection regulatory framework will be welcomed in order to force these and other systems to comply with clear rules regarding the points mentioned and other data subject's rights, as well, as generating trust between the several actors involved.



