Proteção de Dados e Inteligência Artificial: Perspectivas Éticas e Regulatórias

The Use of Big Data in Education Policy[1]

O Uso do Big Data nas Políticas Púbicas Educacionais

CLARICE SEIXAS DUARTE[2]

PhD in Philosophy and General Theory of Law, Faculty of Law, University of São Paulo, Professor of the Stricto Sensu Graduate Program in Political and Economic Law at Mackenzie Presbyterian University, Coordinator of the Internationalization Project "Social Inclusion, Public Policy and Governance to Reduce Inequalities" (CAPES PRINT), Representative in Brazil of the International Platform "Academics Stand Against Poverty" (ASAP), acting as scientific coordinator of the right to education nucleus.

FAYOLA CAJUELLA

Master of Law and Development, Escola de Direito de São Paulo da Fundação Getulio Vargas (FGV Direito SP), Bachelor of Law, Faculdade de Direito de São Bernardo do Campo, Practicing Lawyer in São Paulo. Researcher at FGV.

FERNANDO BÜSCHER VON TESCHENHAUSEN EBERLIN[3]

Doutorando e Mestre em Direito Político e Econômico pela Universidade Mackenzie, MBA em Ciências do Consumo Aplicadas pela Escola Superior de Propaganda e Marketing, Especialista em Direito da Responsabilidade Civil pela FGV Direito SP. Foi pesquisador visitante na Faculdade de Direito da Universidade de Cambridge. Professor nos cursos de pós-graduação *lato sensu* da FGV Direito SP, Advogado em São Paulo.

ABSTRACT: The article analyzes if big data can be used in education policy, comparing this tool with the mere use of official statistics. By using a bibliographic and documentary methodology, the paper makes a revision of overseas literature regarding the benefits arising from the use of big data in public policies and addresses the risks involving its use, particularly those associated with privacy and the potential misunderstanding of information gathered by means other than traditional methods (like censuses and scientific method-based research). The article also analyses official documents such as constitutional and ordinary norms referring to the right of education in order to check how the Brazilian National Education Plan (Brazilian Law 13.005/2014) drives education policy in connection with the set of data and how big data might integrate with traditional information-gathering methods. The paper concludes that rules and assurances must be created to enable the legal safety of data-treatment operations, as well as the protection of children and teenagers as users.

KEYWORDS: Big data; personal data protection; education policy; Brazilian National Education System; Brazilian National Education Plan.

RESUMO: O artigo investiga se o *big data* pode ser utilizado nas políticas públicas educacionais, comparando essa ferramenta com o uso puro e simples das estatísticas oficiais. Utilizando uma metodologia bibliográfica e documental, é realizada a verificação da posição doutrinária, em especial fora do Brasil, a respeito das vantagens do uso dessa ferramenta no campo das políticas públicas. Também são estudados os riscos do uso do *big data*, em especial os relativos à privacidade e à possibilidade de má compreensão das informações colhidas por outros mecanismos que não decorram de métodos tradicionais, como os censos e as pesquisas baseadas em metodologias científicas. Por meio da análise de documentos oficiais, como normas constitucionais e infraconstitucionais relativas ao direito à educação, o trabalho verifica como o Plano Nacional de Educação orienta as políticas educacionais ao uso de dados e como o *big data* poderia se integrar aos métodos tradicionais de obtenção de informações. Ao final, conclui-se que é necessário criar regras e garantias que viabilizem a segurança jurídica de operações de tratamento de dados, bem como a proteção das crianças e adolescentes enquanto usuários.

PALAVRAS -CHAVE: *Big data;* proteção de dados pessoais; políticas públicas educacionais; Sistema Nacional de Educação; Plano Nacional de Educação.

INTRODUCTION

Public policies are government programs stemming from a set of processes intended to achieve previously and politically defined objectives, selecting priorities, setting aside the required resources, and setting goals to ensure attainment of the expected outcomes. The Brazilian Data Protection Act (Law 13.709/2018, or "Lei Geral de Proteção de Dados Pessoais" – LGPD) authorizes public entities to process personal data (including sensitive data) in order to implement public policies (art. 7, III and art. 11, II, b)[4]. In the education area, Brazilian Law 13.005/2014, which creates the National Education Plan ("Plano Nacional de Educação" – PNE), provides the parameters required to draft education policies and indicates the importance of data gathering for this purpose. Brazilian Law 12.965/2014 (the Brazilian Civil Rights Framework for the Internet, or "Marco Civil da Internet" – MCI), in its turn, acknowledges the importance of the World Wide Web in the education arena as a tool for exercising citizenship and promoting culture (art. 26).

One of the contemporary world's characteristics is the large amount of information available on the Internet, coming from all manner of sources (social networks, search engines, personal information, corporate information, etc.).

---

4   This law, as explained by Mendes and Doneda has 5 main axes which are: i) unity and generality of law enforcement; ii) legitimation for data processing by the express provision of the legal hypotheses; iii) principles and rights of the data subject; iv) obligations of data processing agents; and v) accountability of agents (MENDES, Laura Schertel; DONEDA, Danilo. Reflexões iniciais sobre a nova Lei Geral de Proteção de Dados. *Revista de Direito do Consumidor*, São Paulo, v. 120, p. 469-483, nov. 2018. p. 471).

This dataset, also known as big data, can be used in the public policy field to supplement the official statistics ordinarily used by government bodies. So, the research problem of this paper consists in investigate whether the use of big data is possible and adequate for the implementation of PNE, considering that its goals (and the strategies to achieve them) are complex, being the use of information essential for this purpose. Therefore, this paper aims to discuss if big data can help to improve the outcomes (e.g. education quality and access) of existing public policies. The methodology adopted is essentially bibliographic and documentary, by consulting academic articles on the subject of big data and public policies, as well as official documents such as constitutional and ordinary norms referring to the right to education. In addition, it was considered quantitative data related to PNE goals and education in Brazil, available in reports elaborated by organizations such as the PNE Observatory[5] and the Anisio Teixeira National Institute of Educational Studies and Research[6]. The article uses the hypothetical-deductive method, which allows an analysis that, according to Marconi and Lakatos[7], consists – from the perspective of Popper's studies[8] – to perceive problems, gaps or contradictions in the previous knowledge or existing theories. Considering the obstacles for the achievement of PNE's objectives, this paper proposes to discuss the relevance of tools (e.g. big data) to fill the gaps for diagnosis and decision making in educational public policies. Thus, the proposed discussion also justifies the Law and Public Policy approach (developed by both Bucci and Coutinho[9] as well as Bucci and Duarte[10]), which is a methodology that allows examination and a better understanding of complex and large in scale juridical problems, whose solutions require an interdisciplinary dialogue. In this paper, the legal (and large in scale) problem regards PNE objectives implementation, which involves the interdisciplinary dialogue among Law, Technology and Educational Policy issues.

With regard to the structure of this paper, Section 1 reviews the existing literature on the concept of big data and the current state of art concerning the use of this tool in the public policy domain. In order to address the discussions about the criteria for big data use and the privacy harms it may cause, Section 2 investigates the benefits and hazards of governmental use of information

---

5   OBSERVATÓRIO DO PNE. Metas do PNE. Available at http://www.observatoriodopne.org.br/metas-pne. Last viewed on Dec.12.2017.

6   INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANISIO TEIXEIRA. *Resumo Técnico*: Censo da Educação Básica 2018. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira, 2019.

7   MARCONI, M. de A.; LAKATOS, E. M. *Metodologia científica*. 5. ed. 4. São Paulo: Atlas, 2010, p. 73.

8   POPPER, K. R. *Conhecimento objetivo*: uma abordagem evolucionária. São Paulo: Itatiaia: Edusp, 1975.

9   BUCCI, Maria Paula Dallari; COUTINHO, Diogo Rosenthal. Arranjos jurídico-institucionais da política de inovação tecnológica: uma análise baseada na abordagem de direito e políticas públicas, p. 313-340. In *Inovação no Brasil*: avanços e desafios jurídicos e institucionais. São Paulo: Blucher, 2017.

10   BUCCI, Maria Paula Dallari; DUARTE, Clarice Seixas. *Judicialização da Saúde*: a visão do Poder Executivo. São Paulo: Saraiva, 2017.

from the internet and other bases. Finally, this paper will deepen the analysis of education policy in Brazil, studying how PNE goals can best be achieved using big data tools in their strategies. This will be done in Section 3, which will analyze practical situations in which the use of this kind of information may be viable for improving education policy.

1 BIG DATA AND THE IMPORTANCE OF INFORMATION FOR PUBLIC POLICY

As Bucci notes, public policy must aim to accomplish defined objectives, expressing selected priorities, setting aside the means necessary to achieve them, and the timeframe within which one expects results to be attained[11]. Defining public policies' priorities depends on the accurate understanding of the context and of needs of the people for whom it is intended, in addition to other factors associated with the State role.

Fernández and Ferrer point out that public administration is not an independent variable, but rather one fully dependent on other factors like the economy, politics, technology and social evolution[12]. To draft public policy, government needs data and information on society, as well as its needs and priorities[13] and, in order to get it, the government ordinarily uses official databases[14]. In the information society, the enormous amount of data available may also be helpful.

Given this context, the term big data relates to a diverse and vast set of information that can be defined as databases so massive that current technologies are so far unable to handle appropriately[15]. In addition, and this may be the more important meaning of the term, big data represents the potential for interconnection and relationships among this data. For Williamson, data processing (and not just access to it) provides a "digital snapshot" that reveals more than the mere agglutination of all the information at hand (according to

---

11 BUCCI, Maria Paula Dallari. O conceito de Políticas Públicas em Direito. In BUCCI, Maria Paula Dallari (ed.). *Políticas Públicas*: reflexões sobre o conceito jurídico. São Paulo: Editora Saraiva, 2006, p. 39. Our translation.

12 FERNÁNDEZ, Yarina Amoroso; FERRER, Dévorah Costales. Big Data: una herramienta para la administración pública. *Ciencias de La Información*, [s.l.], v. 47, n. 3, p. 3-8, 2016. p. 6.

13 Fernández and Ferrer stress the need for data management and analysis to obtain social trends, behaviors and statistics as a means to develop social and budget policies based on the wants and demands of citizens, who also must play an active role in the drafting of public policy intended for social sectors (FERNÁNDEZ, Yarina Amoroso; FERRER, Dévorah Costales. Big Data: una herramienta para la administración pública. *Ciencias de La Información*, [s.l.], v. 47, n. 3, p. 3-8, 2016. p. 5.)

14 The 2018 basic education census, for example, provides information for public policies by measuring the number of students enrolled, percentage of students by color/race, number of full time students, as well as teacher and school data (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANISIO TEIXEIRA. *Resumo Técnico:* Censo da Educação Básica 2018. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira, 2019.).

15 STOUGH, Roger; MCBRIDE, Dennis. Big Data and U.S. Public Policy. *Review of Policy Research*, [s.l.], v. 31, n. 4, p. 339-342, 2014. p. 339.

the author, what matters is not to have a large amount of data, but what one does with it)[16].

Big data is characterized by the "4 V's": volume, velocity, variety and veracity[17]. These features lead to a change in the classic statistics-development paradigms. The new paradigm is materialized in the potential to quickly (almost instantly) capture and treat a large volume of data[18], often at a lower cost.

Indeed, as Jarmin and O'Hara note, traditional statistics are specially devised for certain purposes, limited to a universe framed by contents emerging from pre-defined and specific questions, and, in many cases, lack information volume, velocity and even veracity[19]. Hackl adds that certain surveys (like the census) provide results for small geographic areas at a high cost[20] and suggests that the use of big data, in addition to expanding analytical possibilities, may provide efficient results at a lower cost for public administration.

Schintler and Kulkarni indicate that big data is a helpful source for countless aspects of public policy drafting, such as the problem definition, the evaluation of existing policies and even the engagement of citizens and other agents involved in the process[21]. As an example, it can be mentioned that working with data allows surveillance strategies to identify potential terrorist behaviors or other security menaces. It also allows identifying epidemics (e.g. the greater the web searches for flu medications in a certain region, the greater the indication of a potential epidemic and the better the odds of taking preventive steps in the affected area[22]).

Data that might be useful to State action may come from a wide range of sources and even from procedures already incorporated into the economic

---

16 WILLIAMSON, Andy. Big Data and the Implications for Government. *Legal Information Management,* [s.l.], v. 14, n. 04, p. 253-257, Dec. 2014. Cambridge University Press (CUP). http://dx.doi.org/10.1017/s1472669614000553. p. 253.

17 JARMIN, Ron S.; O'HARA, Amy B. Big Data and the Transformation of Public Policy Analysis. *Journal Of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p. 715-721, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21925. p. 715.

18 "The classic statistical paradigm was one in which researchers formulated a hypothesis, identified a population frame, designed a survey and a sampling technique, and then analyzed the results (Groves, 2011). The new paradigm means that it is now possible to digitally capture, semantically reconcile, aggregate, and correlate." (LANE, Julia. Big data for public policy: the quadruple helix. *Journal Of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p. 708-715, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21921. p. 709.)

19 JARMIN, Ron S.; O'HARA, Amy B.. Big Data and the Transformation of Public Policy Analysis. *Journal of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p. 715-721, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21925. p. 716.

20 HACKL, Peter. Big Data: What can official statistics expect? *Statistical Journal of The Iaos,* [s.l.], v. 32, n. 1, p. 43-52, Feb. 27, 2016. IOS Press. http://dx.doi.org/10.3233/sji-160965. p. 43.

21 SCHINTLER, Laurie A.; KULKARNI, Rajendra. Big Data for Public Analysis: the Good, the Bad, and the Ugly. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 343-348, 2014. p. 343.

22 An experiment in this sense was "Google flu trends", whose results can be found at https://www.google.org/flutrends/about/ (last viewed on Apr. 19.2018).

routine. One interesting example that Hackl provides is the use of scanner data (like those used by retailers to scan merchandise barcodes). This information, if provided to the State, would allow quick and effective understanding of the evolution of prices and of food and beverage sales[23]. This would be helpful, for example, to effectively measure inflation rates at a lower cost.

Data-based technology can also promote human development and better life quality. Benkler provides some examples in this field, such as the gains to agriculture from biotechnology innovation (and the potential increase in societies' ability to become self-reliant on nutritious foods); the positive consequences of pharmaceutical research for medicine; and the improvement of education due to the better access to information and materials (such as textbooks)[24].

In the specific area of education, data coming from the various exams applied to students (such as the National High School Exam) may allow the State do verify students' performance and analyze the outcome of public policies, as well as think of new polices that can directly reach the identified problems and produce faster and more consistent results.

In this same line, observing what people say on social networks may be useful to identify a school that is experiencing deficiencies. Indeed, online comments and posts, including those on social media, are resources of information that may help to understand people's opinions and concerns. This information, as Wang points out, is of great value for public policy drafting insofar as it allows understanding the community's problems and needs, as well as drafting policies to address them, evaluating the effectiveness of those policies and getting people engaged in ideas-generating and problem-solving processes[25].

Furthermore, big data can also be a helpful tool to promote public policy transparency and credibility (which are essential to Democracy), generating trust between the various groups of actors – sometimes with different interests – involved in certain discussions[26].

---

23   HACKL, Peter. Big Data: What can official statistics expect? *Statistical Journal of the Iaos,* [s.l.], v. 32, n. 1, p. 43-52, Feb. 27, 2016. IOS Press. http://dx.doi.org/10.3233/sji-160965. p. 44.

24   BENKLER, Yochai. *The Wealth of Networks:* How Social Production Transforms Markets and Freedom. New Haven: Yale University Press, 2006. p. 310.

25   WANG, Yinying. Big Opportunities and Big Concerns of Big Data in Education. *Techtrends,* [s.l.], v. 60, n. 4, p. 381-384, Apr. 27, 2016. Springer Nature. http://dx.doi.org/10.1007/s11528-016-0072-1. p. 382. With reagard to the population's public-policy engagement through big data, it is important to emphasize that the MCI provides certain guidelines for the public administration's use of the Internet (Art. 25), including "strengthening social public-policy participation" (our translation).

26   SCHINTLER, Laurie A.; KULKARNI, Rajendra. Big Data for Public Analysis: the Good, the Bad, and the Ugly. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 343-348, 2014. p. 345.

Data has rather high potential in the public policy arena, but the use of information is not an entirely novel strategy. Indeed, although data is currently more widely available, it has always been used by the public and private sectors (respectively for policy and strategy drafting). The difference is that, in the past, data was directly created by business firms or public entities, which kept and used them internally. Now, with the Internet, data is more freely available and, as mentioned before, is created by individuals themselves on platforms like Facebook, YouTube and Twitter[27].

The novelty of the present times is that data from business firms, agencies, states, municipalities, universities and other information providers can mutually supplement each other and, when integrated, they can form a complete database without the gaps that individual databases may have[28]. Jarmin and O'Hara argue that compiling diverse data sources, particularly when using big data tools, may create an unprecedented opportunity to transform public policy analysis[29]. In the education field, this integration is needed to promote the articulation of all organs dealing with public policies, which is currently a challenge.

As explained above, LGPD (Brazilian Data Protection Act) has already allowed the treatment of personal data (including sensitive data) by public administration for the purposes of treatment and shared use as required for the execution of public policies governed by laws or regulations. Chapter IV of LGPD also includes certain rules that the public administration must abide by while treating personal data, such as using such data for public purposes (art. 23) and adopting personal data-protection principles (art. 26). When dealing with educational policies, these rules must be compliant as well with art. 14 of LGPD, which bring the rules for children and adolescents' personal data processing[30].

Notwithstanding the clear potential for the use of big data in the public policy domain, some risks exist that require caution when using this tool, as analyzed in the next Section.

---

27 WILLIAMSON, Andy. Big Data and the Implications for Government. *Legal Information Management*, [s.l.], v. 14, n. 04, p. 253-257, Dec. 2014. Cambridge University Press (CUP). http://dx.doi.org/10.1017/s1472669614000553. p. 254.

28 LANE, Julia. Big data for public policy: the quadruple helix. *Journal Of Policy Analysis And Management*, [s.l.], v. 35, n. 3, p. 708-715, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21921. p. 710.

29 JARMIN, Ron S.; O'HARA, Amy B.. Big Data and the Transformation of Public Policy Analysis. *Journal of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p. 715-721, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21925. p. 715.

30 In the case of children, LGPD determines that parental consent must be obtained, which can be an issue in Brazil, because of the high rates of illiterates adults (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (Brasil). *Conheça o Brasil – População:* Educação. [2019]. Available at: https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18317-educacao.html. Last viewd on Oct. 18.2019.)

2 POTENTIAL NEGATIVE ASPECTS OF THE USE OF BIG DATA FOR PUBLIC POLICY PURPOSES

Some issues must be addressed before big data can in fact become a part of public policy. Hackl includes among these issues the availability of information technology methods, the implementation of statistical methods that can understand and integrate big data information, the presence of data management, protection policies and guidelines, as well as the existence of laws allowing access to and use of data for these purposes[31].

In a similar vein, Lane indicates four groups of issues that must be addressed for data treatment purposes: (i) technical (determining the provenance of data and aggregating information in scientific manner); (ii) legal (e.g. agreements or regulation allowing information sharing); (iii) privacy (e.g. consent and data anonymization); and (iv) training (building the skills of the people who will be working with the data)[32].

The concerns listed by the foregoing authors may be addressed in two major groups. One relates to how the data is treated, involving the aspects associated with technical, training and big data statistical methods integration issues. The other is linked to the topic of the privacy of the people whose personal information will undergo treatment, which involves the presence of a legal framework authorizing such treatment, as well as information and consent methods (which, according art. 14 of LGPD, are more specific in the case of children and adolescents) required before processing personal data for public policy purposes. This Section's purpose is to address these two aspects.

2.1 THE NEED FOR CRITERIA IN DATA ANALYSIS AND TREATMENT

Inappropriate data treatment may lead to incorrect inferences from reality, enable misunderstandings, and even allow information to be manipulated. As noted earlier, big data is a large information set that can only be of use when this information is correctly processed and understood. As Hackl argues, one cannot say that big data is capable of completely replacing official statistics in the public policy domain[33]. Indeed, statistics are based on scientific methodologies and assumptions that must be preserved in order to make a proper interpretation of the information.

Stough and McBride note that the characteristics of big data include the irregularity and heterogeneity of the information available, and that current

---

31 HACKL, Peter. Big Data: What can official statistics expect? *Statistical Journal Of The Iaos,* [s.l.], v. 32, n. 1, p. 43-52, Feb. 27, 2016. IOS Press. http://dx.doi.org/10.3233/sji-160965. p. 47.

32 LANE, Julia. Big Data for Public Policy: the Quadruple Helix. *Journal Of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p.708-715, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21921. p. 712.

33 HACKL, Peter. Big Data: What can official statistics expect? *Statistical Journal Of The Iaos,* [s.l.], v. 32, n. 1, p.43-52, Feb. 27, 2016. IOS Press. http://dx.doi.org/10.3233/sji-160965. p. 52.

statistical analysis methods are not yet prepared to properly analyze such data[34]. This assumption challenges the integration of big data with official statistics methodologies, but it is important to address this issue, once state action, according to Amorim and Sacaff[35], requires planning and the data used must be consistent, even if they are heterogeneous.

In order to address these issues, one of the first precautions to be taken is to identify potential deviations and biases of the information gathered. Bright and Margetts point out that when it comes to big data, even if the data are readily available, they may be biased and unrepresentative; in addition, despite low-cost availability, there may be limitations (legal and contractual) on their use, and the data may also be subject to fraud and distortions[36]. In this sense, Schintler and Kulkarni warn that the use of big data as a mere set of numbers or words without taking account of time, spatial, social and cultural aspects may lead to false interpretations and conclusions[37].

Bright and Margetts, in a specific analysis of public policies, bring up an important characteristic of big data, which is the fact that the information gathered from social media is "passively" provided. Indeed, people on social media may express their views about politics and the government in a carefree way, without expecting this information to be collected and used in public policy drafting process. According to the authors, such passive contribution is very different from mechanisms by means of which citizens actively deliberate and take part in forums, public hearings and discussions; in these scenarios, people tend to make deeper reflections about the proposed policies in order to make conscious choices in coherence with public policy drafting[38].

Therefore, big data's characteristics suggest a constant need to analyze the quality of the data in use, which raises questions regarding certain automatic decision-making mechanisms[39].

---

34 STOUGH, Roger; MCBRIDE, Dennis. Big Data and U.S. Public Policy. *Review Of Policy Research,* [s.l.], v. 31, n. 4, p. 339-342, 2014. p. 339.

35 AMORIM, M. D.; SCAFF, E. A. S.. O planejamento como instrumento de gestão da educação básica: tendências das políticas atuais. In: Antonio Bosco de Lima; Dirce Nei Teixeira de Freitas. (Org.). *Políticas sociais e educacionais*: cenários e gestão. 1ed.Uberlandia, MG: EDUFU, 2013, v. 1, p. 247-266.

36 BRIGHT, Jonathan; MARGETTS, Helen. Big Data and Public Policy: Can It Succeed Where E-Participation Has Failed? *Policy & Internet,* [s.l.], v. 8, n. 3, p. 218-224, Sep. 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/poi3.130. p. 221.

37 SCHINTLER, Laurie A.; KULKARNI, Rajendra. Big Data for Public Analysis: the Good, the Bad, and the Ugly. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 343-348, 2014. p. 344.

38 BRIGHT, Jonathan; MARGETTS, Helen. Big Data and Public Policy: Can It Succeed Where E-Participation Has Failed? *Policy & Internet,* [s.l.], v. 8, n. 3, p. 218-224, Sep. 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/poi3.130. p. 222.

39 Indeed, one of the features inherent to data treatment is the use of algorithms, pre-programmed formulas to interpret the information gleaned from the Internet, detect patterns and predict behaviors. Algorithms may operate with or without human interference. In the latter case, there are many questions involving the potential to submit personal data do automated decision-making mechanisms. In Europe, for example, Article 21 of regulation 2016/679 (Available at: http://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0

Athey explains that big data offers several means to predict behaviors and events and that one of these techniques, called "self machine learning" (SML), consists in the use of data and information showing repetitive patterns to forecast certain situations (e.g. student absences that must be detected and reported to the guardian council). Such repetitive information can be processed by machines using algorithms and making important decisions with no human interference. For example, the public administration might, without human interference, determine how to deploy police officers based on an algorithm using information on the number of break-ins, robberies and other incidents, population density data, and even Google Street View images. In this respect, Athey also warns that, although this kind of technology may be extremely efficient, the algorithms must be correctly programmed to prevent use of misguided or out-of-context assumptions (such as deeming a certain pattern repetitive when, in fact, it is not). This is particularly important in cities showing heterogeneous behavior patterns[40].

Despite the reduced impact when algorithms are properly set up, the most prudent way to use big data for public policy purposes appears to be maintaining some kind of human intervention in decision-making. In other words, policy makers shouldn't simply migrate to algorithms and automated decisions processes, but they should improve the current state of art, by using the potentialities and benefits of big data together with the existing methods.

2.2 Privacy of the citizens whose data the government uses to draft public policy

Potential privacy breaches also require attention while using big data in association with public policy. The concept of privacy in today's information society has certain vicissitudes. In addition to meaning the impossibility of third--party interference in aspects of personal life, privacy in the data society also means giving the owner of the personal data the power to determine what will be done with the information available on them (whether or not it will be used, disclosed, treated)[41]. For this reason, data protection laws state that this type of information can only be processed when there is legal authorization and when the purpose of data collection is observed.

679&from=PT. Last viewed on Nov. 08.2017.), allows the owner of data to oppose the submission of their personal information to automated decision-making mechanisms. However, exceptions provided for the use of data under certain circumstances, including "pursuit of attributes in public interest," which spices up the discussion of the legitimate use of personal data in the public policy domain.

40 ATHEY, Susan. Beyond prediction: Using big data for policy problems. *Science,* [s.l.], v. 355, n. 6324, p.483-485, Feb 2, 2017. American Association for the Advancement of Science (AAAS). http://dx.doi.org/10.1126/science.aal4321. p. 1.

41 "[...] tutelage of privacy, under this new interpretation, is not limited to a ban on foreign intrusion into intimate life (general duty to abstain). It also imposes positive duties such as the duty to request consent to include a certain person's name in a data file, or the duty to enable the owner of the data to correct entries in such files at any time." (SCHREIBER, Anderson. *Direitos da Personalidade*. São Paulo: Atlas, 2011. p. 131.)

The topic becomes more complex in the field of government policy purposes, since the collection of personal data without knowing what will be done may represent a waste of public resources (since the processing of data involves investments in sophisticated technologies), besides harming the privacy of citizens. Stough and McBride warn that the use of big data for public policy purposes may represent rather relevant fears that the government will explore and trespass upon people's privacy sphere[42]. Rodotà states that the risks arising from surveillance society are normally linked to the political use of information to control citizens' behavior, which may lead to dictatorial societies[43]. Tomasevicius Filho brings up privacy-breach attempts at the initiative of European governments, particularly in France and Germany[44]. The concern is justified because public interest is one of the elements that authorize the use of personal data in the European Community, including in an automated manner, regardless of the owner's consent[45].

Another important privacy-related aspect is the tendency, particularly in the public sector, for the data to be open and transparent, which, according to Williamson, is an important element for democracy[46]. Rodotà, following this argument, states that the most democratic systems give great value to transparency, while dictatorial logic is based in rules of secrecy and opacity[47]. Openness usually materializes in open-data policies, a trend found in several

---

42 "[...] the biggest impediment to Big Data or information concerning people is that its use potentially intrudes into or supports the perception that it might intrude into and compromise our constitutionally granted protection of privacy from government exploitation." STOUGH, Roger; MCBRIDE, Dennis. Big Data and U.S. Public Policy. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 339-342, 2014. p. 341.

43 RODOTÀ, Stefano. *A vida na sociedade da vigilância – a privacidade hoje*. Tradução: Danilo Doneda e Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008. p. 113.

44 "There was an attempt in France to create the SAFARI (Automated Individual Information and Administrative Files System), which was terminated in 1974 after complaints. In Germany, in 1982, the population was expected to answer 160 questions in a census. It was understood that a deviation from purpose existed in this case, where people had no control over what would be done with this information. The German courts, while examining the case, mentioned informational self-determination as the individual right to decide the limits to which their personal data could be used." (TOMASEVICIUS FILHO, Eduardo. Em direção a um novo 1984? A tutela da vida privada entre a invasão de privacidade e a privacidade renunciada. *Revista da Faculdade de Direito da Universidade de São Paulo,* [s.l.], v. 109, p. 129-169, 2014. p. 152. Our translation.)

45 Article 6, item 1, sub-item e of Regulation 2016/679 (available at: http://eur-lex.europa.eu/legal-content/PT/TXT /PDF/?uri=CELEX:32016R0679&from=PT. Last viewed on Nov.08.2017.) includes a general rule according to which the treatment of personal data is legal even in the absence of the owner's consent if needed to perform functions in public interest or exercise a public authority invested in the official responsible for such treatment. Article 89 of the same Regulation provides rules for the use of data in public interest, such as the Member States' obligation to ensure the data owner's rights and liberties, guarantee the principle of data minimization (that is, to threat only information as strictly necessary for the purposes of public interest, and embrace data pseudonymization or anonymization measures wherever possible.

46 WILLIAMSON, Andy. Big Data and the Implications for Government. *Legal Information Management*, [s.l.], v. 14, n. 04, p. 253-257, Dec. 2014. Cambridge University Press (CUP). http://dx.doi.org/10.1017/s1472669614000553. p. 254.

47 RODOTÀ, Stefano. *A vida na sociedade da vigilância – a privacidade hoje*. Translation: Danilo Doneda and Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008. p. 45.

countries, including Brazil, as the www.dados.gov.br website shows[48]. Open data policies are based on the notion that certain kinds of information must be available, regardless of intellectual property rights, patents, censorship, or any other constraints that normally apply to data[49]; in the government context, open data policies create alternatives for individuals, for the private sector and for the profit and nonprofit organizations that may find new opportunities and develop new products and services[50].

Data openness is an additional element reinforcing the need for control and appropriateness of the procedures involving the use of personal information, so as to prevent breaching of privacy. One must measure the benefits stemming from the use of personal data in order to determine acceptable limits for its use[51].

To ensure citizens' rights against State surveillance, Rodotà suggests that data protection regulations should forbid generalized interconnections between data, keeping governmental data sources disaggregated, except by situations expressly allowed by law[52]. In this same line, Simitis understands that government cannot be understood as a unique information unit allowing unlimited data transfers between its entities; governmental structures should be rearranged to create a functional segmentation concerning the use of data[53].

Despite these positions, it seems reasonable to argue that the interconnection of data between government entities could be done in specific situations involving a relevant public interest (e.g. the evaluation of educational public policies), since this public interest is balanced with the individual guarantees stemming from the right to privacy.

---

48   Brazilian Law 12.527 ("Lei de Acesso à Informação" – LAI) establishes that the State must provide access to information by means that are objective, fast, transparent, clear and of easy understanding.

49   Open Data Policies aim to encourage transparency and openness in the private and specially in the public sector; these policies may involve many categories of data, such as information concerning cultural data (titles and authors), scientific information (produced in scientific research – astronomy, zoology etc.), financial indicators (government accounts and information on financial markets), statistics (census and socioeconomic indicators), weather (information to predict and understand climate) and environment (level of pollutants, quality of water) (CANNATACI, Joseph A. *Report of the Special Rapporteur on the right to privacy.* New York: United Nations, 2017. Available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&c d=1&ved=0ahUKEwi50pHB4sPZAhUEPVAKHXOnBdsQFgguMAA&url=http%3A%2F%2Fwww.ohchr.org% 2FDocuments%2FIssues%2FPrivacy%2FA-72-43103_EN.docx&usg=AOvVaw0Ekn-qUgeqxHsOEU2XMinv; last viewed on Apr.19.2018. p. 16.).

50   CARLO, Bertot John et al. Big data, open government and e-government: Issues, policies and recommendations. *Information Polity,* [s.l.], v. 19, n. 12, p.5-16, 2014. IOS Press. http://dx.doi.org/10.3233/IP-140328. p. 6.

51   In this respect, Schintler and Kulkarni use cost-benefit analyses to discuss how to measure an acceptable breach of the privacy of a group of people if such a breach would enable implementing a public policy that will benefit an entire certain region (SCHINTLER, Laurie A.; KULKARNI, Rajendra. Big Data for Public Analysis: the Good, the Bad, and the Ugly. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 343-348, 2014. p. 346.).

52   RODOTÀ, Stefano. *A vida na sociedade da vigilância* – a privacidade hoje. Translation: Danilo Doneda and Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008. p. 175.

53   SIMITIS, Spiros. Reviewing Privacy in an Infomation Society. *Univsersity Of Pennsylvania Law Review,* [s.l.], v. 135, n. 3, p.707-746, 1987. Available at: http://scholarship.law.upenn.edu/penn_law_review/vol135/ iss3/3. Last viewed on Jan. 05. 2018. p. 741.

A very relevant suggestion on this front would be to create mechanisms such as rules imposing that data used for public policy purposes should be anonymized or submitted to some kind of privacy enhancing technology[54]. In other words, personal information enabling individual identification must be deleted or transformed before being used for the purposes of analysis[55]. Specific cases in which information must be associated with a certain person must be treated as exceptions.

3 BIG DATA IN EDUCATION POLICY

Wang notes that students today are growing up in a digital environment, where they carry out their daily activities as well as learning tasks, leaving behind abundant digital traces like clicks, text typed into Websites, tablets and smartphones, sensor-detected movements, and facial expressions recognized by ubiquitous cameras[56]. This creates a huge database that allows a better understanding and knowledge of students, their needs, yearnings, orientations and deficiencies, but also allows this information to generate trails that enable business firms to track their users' behavior.

In order to protect children, the Unicef report "The State of World's children 2017. Children in a digital word" recommends priority actions in benefit of children and to protect the most vulnerable among them, as follows: (i) to safeguard the privacy and identity of children online; (ii) to keep them safe from online violence (abuse, exploitation and other hazards); (iii) to provide guidance in connection with potential threats and a safe browsing stance; and (iv) to foster the private sector's ethical action as concerns standards and practices for the protection of children connected to the world wide web[57].

Such measures, together with correct use of information, may improve learning quality. Indeed, as Schintler and Kulkarni point out, public policy is a response to a current, perceived or anticipated problem[58], so that, in the education area, knowledge of the data would help understand a given school's deficiencies, improve learning quality wherever needed, identify locations where

---

54 As Doneda points out, privacy enhancing technologies can promote changes in the architecture of data processing systems and constitute not only a series of isolated private actions, but also an instrument for public policies that encourage the adoption of these technologies from the conception of products and services (DONEDA, Danilo. *Da Privacidade à Proteção de Dados Pessoais.* Rio de Janeiro: Renovar, 2006. p. 369.). This rationale can also be applied to the policies mentioned in this paper.

55 STOUGH, Roger; MCBRIDE, Dennis. Big Data and U.S. Public Policy. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 339-342, 2014. p. 342.

56 WANG, Yinying. Big Opportunities and Big Concerns of Big Data in Education. *Techtrends,* [s.l.], v. 60, n. 4, p. 381-384, Apr. 27, 2016. Springer Nature. http://dx.doi.org/10.1007/s11528-016-0072-1. p. 382.

57 UNICEF. The State of the world's children 2017. Children in a digital word. Available at https://www.unicef.org/publications/files/SOWC_2017_ENG_WEB.pdf. Last viewed on Dec.12.2017. p. 5.

58 SCHINTLER, Laurie A.; KULKARNI, Rajendra. Big Data for Public Analysis: The Good, The Bad, and The Ugly. *Review of Policy Research,* [s.l.], v. 31, n. 4, p.343-348, 2014. p. 344.

more schools are needed (due to the overpopulation of existing ones), determine what schools need more financial resources, and several other measures.

The PNE recognizes the importance of information for public policy drafting. According to Article 4 of Law 13.005/2014, for example, the PNE's goals should use the National Household Sample Survey ("Pesquisa Nacional por Amostra de Domicílio" – PNAD) for reference purposes, in addition to the demographic census and the national basic and higher education censuses.

Regular tracking of the PNE's implementation and attainment of its goals are carried out by: the Ministry of Education ("Ministério da Educação" – MEC); the House Education Committee ("Comissão de Educação da Câmara dos Deputados") and the Senate Education, Culture and Sports Committee ("Comissão de Educação, Cultura e Esporte do Senado Federal"); the National Education Board ("Conselho Nacional de Educação" – CNE); and the National Education Forum ("Fórum Nacional de Educação"), pursuant to the PNE, Article 5.

For the purposes of goal attainment, Article 7 of the PNE requires the collaborative action of the Union, States, Federal District and Municipalities, which are to create mechanisms for regional monitoring towards attainment of the PNE's goals and those of the respective local plans, so as to produce cooperation among the Federative Entities, adding to the viability of supplementation via national and local coordination and reciprocal cooperation mechanisms.

The common denominator to the plans of each Federative Entity lies based on the following strategies, according to PNE Article 8, Paragraph 1: to ensure the articulation of education policies with other social policies; attention to the unique traits of rural, indigenous and *Quilombola* populations, assuring educational equanimity and cultural diversity; precautions in connection with special education in an inclusive context at all levels, phases and modes and fostering of inter-federative articulation for the purposes of education policy execution. For the purposes of Plan drafting and adjustment, Big Data can be used across Federative Entities with the cooperation of representatives from the teaching community and other members of society.

Article 11, Paragraph 1, of the aforementioned mandates that the National Basic Education Evaluation System must produce, every other year, scholarly performance indicators for the performance of students, taking account of evaluations and data gathered by the basic education census. As for vocational training, goal #11 of the PNE – covering middle-school vocational education – strategy 11.14 proposes structuring a national vocational information framework, cross-referencing training institutions' data with the jobs market. The blending of corporate databases, jobs market statistics, and school information (including student data and evaluations) would be a relevant step toward meeting the goals set forth in the laws mentioned above.

The use of digital media for learning purposes is also part of the PNE's scope. Strategy 7.20 of Goal #7, which covers basic education, establishes guidelines for the provision of digital technology equipment and resources for educational purposes. In the same vein, strategies 12.15 and 14.7 suggest institutionalizing digital-archive formation programs. Student access to digital learning methods, including digital archives, creates a database on their behavior, performance, research underway, and other information that may be relevant to public-policy drafting[59].

The points addressed by PNE goal #3 include the requirement that the net rate of enrollment in middle school must rise to 85% by 2024. In order to achieve this goal, strategy 3.9 was drafted as follows: "to promote the active pursuit of the population between the ages of 15 and 17 years old not enrolled in school, in articulation with social service, health-care and teenagers and youths protection services"[60].

Strategies 3.12 and 3.13, too, try to protect teenagers living in contexts that may lead to quitting school. Strategy 3.12 tries to provide choices for middle education, with assured quality, to those whose parents hold jobs of an itinerant nature, while strategy 3.13 prescribes the adoption of policies preceding cases where students drop out due to prejudice or other forms of discrimination, outlining a protection network against exclusion.

An example of the use of basic technology to prevent quitting school can be found in student attendance controls. The municipal student monitoring system works as follows: the student's ID card is run through a barcode scanner after they enter the school premises. The parents or tutors then receive an email or cell-phone (text or app) message noting entrance. Some systems show the school calendar, out-of-class activities, exam corrections and grades availability, which allows a child's parents or legal tutors to directly monitor their progress.

When a student is continuously or repeatedly absent, social workers are designated to look into and report the reasons for this. To this end, they visit the student's home and talk to their parents and tutors to determine the context in which they exist, which allows minimizing difficulties and proposing solutions.

In this sense, social workers can help families obtain school bus passes, enroll in welfare programs, or, as the case may be, refer cases of violence or hazard to the Special Social Service Reference Center ("Centro de Referência Especializado de Assistência Social" – CREAS) and activate the Board of Tutors

---

59 Along these lines, Article 26 of the MCI recognizes the importance of the Internet for Education Policy insofar as it determines that the State, while providing education, must carry out training "for the safe, conscious and responsible use of the Internet as a tool to exercise citizenship, promote culture, and develop technology." Article 27 of the MCI mentions public initiatives to foster the digital culture and "promote the internet as a social tool."

60 Law 13.005/2014, establishing the National Education Plan.

– regarding these issues, teachers often are the key persons to identify cases of child abuse and it is very important to provide them a tool to forward this kind of information. In addition to enabling control on the part of the school and parents, this method, within a big data context, enables a macro outlook for the purposes of tracking goal #3, concerned with controlling dropout rates, and enables effective actions to keep students in school.

The information gathered by social workers may become part of databases maintained by the school, regional education divisions, and local and state Education Bureaus to help form Public Policy propositions or improvements.

It is worth emphasizing that platforms already exist in the Brazilian education arena that gather and use data to track the subject. The Monitoring and Warning Issuance System ("Sistema de Monitoramento e Expedição de Alerta")[61], for example, helps track attainment of PNE goals. The Lemann Foundation, in its turn, has since 2012 maintained the Qedu Platform[62], with basic education system data. The presence of data and data sources to improve learning processes is a given. However, as noted earlier, mere access to the data is not enough for efficient use of Big Data. The data must be worked, and a "data culture" in the realm of education must be formed.

From a practical viewpoint, data, when properly used, can male public policy better reflect people's needs, opinions and yearnings, making the process of policy drafting and implementation more democratic.

Big data can also help build an efficient National Education System. Indeed, given Brazil's heterogeneity and sheer size, lots of municipalities don't have the minimum resources and guidelines to implement a municipal education system. Databases of clear guidelines can enable access to educational system implementation information and guidelines, as well as the provision of information, on the part of such municipalities, to allow assessing – perhaps even automatically – each municipal education system's most important wants and needs.

Data cross-referencing mechanisms also allow obtaining more information on education workers. Data including attendance rate, performance of classes associated with a certain teacher, academic background, professional path (subjects delivered in the past) and many more can not only help to map out a teacher's performance and development, but also can help to define professional training needs and strategies. Working with information systems will also help teachers to understand the students with whom they will be working, besides expediting information gathering for the purposes of reports and other types of

---

61　Available at: https://pne.tce.mg.gov.br/#/public/inicio. Last viewed on Dec. 07.2017.
62　Available at: http://qedu.org.br/. Last viewed on Dec. 07.2017.

time-consuming paperwork. This would allow them to dedicate more time to learning activities, with productivity gains across the system.

Clearly, big data can be used in connection with education policy. However, the risks mentioned in the preceding section may also weigh on the use of the tool in the education arena. According to Wang, the main concerns involving the use of big data in the domain of education lie associated with safety, privacy and ethical issues concerning personal data protection. Safety and privacy interrelate and concern the processes of storage and the appropriate use of data, in addition to required consent to collect and use such data, as the previous Section has discussed. One of Wang's concerns is the inappropriate use of students' lives data (like a deficiency, or a poor performance) in the future, hampering access to universities or even jobs. The data must be protected and used according to the owner's manifest consent. The ethical aspect, in its turn, relates with data owners' knowledge about the use of their information for the purposes of research and policy drafting[63].

CONCLUSION

The use of big data for the purposes of education policy may improve the use of official databases and, besides, bring into the PNE's scope a data – and information-usage culture. This would help produce more accurate diagnoses of regional wants and needs, so that big data might become an important tool for the articulation and integration of a National Education System.

However, education policies using big data must define, clearly and in advance: (i) what data is relevant to education policy; (ii) what data must individually identify students and what can be anonymized; (iii) the purpose for which the data will be collected, preventing inappropriate use of information that may, in the future, harm students' development and admission into academic and professional areas.

In this sense, the form of collection, processing, analysis and – ultimately – disposal of the data must be the focus of public administrators' attention while addressing the lifecycle and the flow of the data obtained. Legal safety and user protection are the cornerstones of the framework for the use of personal data and virtual tools.

To preserve privacy, the use of big data must be driven, whenever possible, by data anonymization, preventing the identification of individual conducts, tastes and interests and focusing on the relevant behavioral aspects of groups of people that may benefit from public policy.

---

63   WANG, Yinying. Big Opportunities and Big Concerns of Big Data in Education. *Techtrends*, [s.l.], v. 60, n. 4, p. 381-384, Apr. 27, 2016. Springer Nature. http://dx.doi.org/10.1007/s11528-016-0072-1. p. 383.

REFERENCES

AMORIM, M. D.; SCAFF, E. A. S. O planejamento como instrumento de gestão da educação básica: tendências das políticas atuais. In: Antonio Bosco de Lima; Dirce Nei Teixeira de Freitas. (Org.). *Políticas sociais e educacionais*: cenários e gestão. 1ed. Uberlandia, MG: EDUFU, 2013, v. 1, p. 247-266.

Associação dos Membros dos Tribunais de Contas do Brasil; Instituto Rui Barbosa. Metas do Plano de Educação (PNE): relatório final do Grupo de Trabalho Atricon-IRB. Cuiabá, 2016. 200p. Available at: https://portal.tce.rs.gov.br/pne/2016/relatorio_atricon_irb/files/assets/common/downloads/Relatorio%20Grupo%20Atricon%20IRB%202016.pdf. Last viewed on Dec. 14. 2017.

ATHEY, Susan. Beyond prediction: Using big data for policy problems. *Science,* [s.l.], v. 355, n. 6324, p. 483-485, 2 fev. 2017. American Association for the Advancement of Science (AAAS). http://dx.doi.org/10.1126/science.aal4321.

BENKLER, Yochai. *The Wealth of Networks:* How Social Production Transforms Markets and Freedom. New Haven: Yale University Press, 2006.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. *DOU*. Brasília, DF, 18 nov. 2011. Lei de Acesso à Informação.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. *DOU*. Brasília, DF, 24 abr. 2014. Marco Civil da Internet.

BRASIL. Lei nº 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. *DOU*. Brasília, DF, 26 jun. 2014.

BRASIL. Lei nº 13.079, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). *DOU*. Brasília, DF, 15 ago. 2018.

BRIGHT, Jonathan; MARGETTS, Helen. Big Data and Public Policy: Can It Succeed Where E-Participation Has Failed? *Policy & Internet,* [s.l.], v. 8, n. 3, p. 218-224, Sep. 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/poi3.130.

BUCCI, Maria Paula Dallari. O conceito de Public Policy em Direito. In BUCCI, Maria Paula Dallari (ed.). *Public Policy:* reflexões sobre o conceito jurídico. São Paulo: Editora Saraiva, 2006.

BUCCI, Maria Paula Dallari; COUTINHO, Diogo Rosenthal. Arranjos jurídico-institucionais da política de inovação tecnológica: uma análise baseada na abordagem de direito e políticas públicas, p. 313-340. In *Inovação no Brasil*: avanços e desafios jurídicos e institucionais. São Paulo: Blucher, 2017.

BUCCI, Maria Paula Dallari; DUARTE, Clarice Seixas. *Judicialização da Saúde*: a visão do Poder Executivo. São Paulo: Saraiva, 2017.

CANNATACI, Joseph A. *Report of the Special Rapporteur on the right to privacy.* New York: United Nations, 2017. Available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwi50pHB4sPZAhUEPVAKHX0nBdsQFgguMAA&url=http%3A%2F%2Fwww.ohchr.org%2FDocuments%2FIssues%2FPrivacy%

2FA-72-43103_EN.docx&usg=AOvVaw0Ekn-qUgeqxHsOEU2XMinv; last viewed on Apr. 19.2018.

CARLO, Bertot John et al. Big data, open government and e-government: Issues, policies and recommendations. *Information Polity,* [s.l.], v. 19, n. 12, p. 5-16, 2014. IOS Press. http://dx.doi.org/10.3233/IP-140328.

DONEDA, Danilo. *Da Privacidade* à *Proteção de Dados Pessoais.* Rio de Janeiro: Renovar, 2006.

FERNÁNDEZ, Yarina Amoroso; FERRER, Dévorah Costales. Big Data: una herramienta para la administración pública. *Ciencias de La Información,* [s.l.], v. 47, n. 3, p. 3-8, 2016.

HACKL, Peter. Big Data: What can official statistics expect? *Statistical Journal Of The Iaos,* [s.l.], v. 32, n. 1, p.43-52, 27 Feb. 2016. IOS Press. http://dx.doi.org/10.3233/sji-160965.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (Brasil). *Conheça o Brasil –* População: Educação. [2019]. Available at: https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18317-educacao.html. Last viewed on Oct. 18.2019.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANISIO TEIXEIRA. *Resumo Técnico:* Censo da Educação Básica 2018. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira, 2019.

JARMIN, Ron S.; O'HARA, Amy B.. Big Data and the Transformation of Public Policy Analysis. *Journal of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p. 715-721, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21925.

LANE, Julia. Big Data for Public Policy: The Quadruple Helix. *Journal of Policy Analysis And Management,* [s.l.], v. 35, n. 3, p. 708-715, May 10, 2016. Wiley-Blackwell. http://dx.doi.org/10.1002/pam.21921.

MARCONI, M. de A.; LAKATOS, E. M. *Metodologia científica.* 5. ed. 4. São Paulo: Atlas, 2010.

MENDES, Laura Schertel; DONEDA, Danilo. Reflexões iniciais sobre a nova Lei Geral de Proteção de Dados. *Revista de Direito do Consumidor,* São Paulo, v. 120, p. 469-483, nov. 2018.

OBSERVATÓRIO DO PNE. Metas do PNE. Available at http://www.observatoriodopne.org.br/metas-pne. Last viewed on Dec. 12.2017.

POPPER, K. R. *Conhecimento objetivo:* uma abordagem evolucionária. São Paulo: Itatiaia: EDUSP, 1975.

RODOTÀ, Stefano. *A vida na sociedade da vigilância –* a privacidade hoje. Translation: Danilo Doneda and Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008.

SCHINTLER, Laurie A.; KULKARNI, Rajendra. Big Data for Public Analysis: the Good, the Bad, and the Ugly. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 343-348, 2014.

SCHREIBER, Anderson. *Direitos da Personalidade.* São Paulo: Atlas, 2011.

SIMITIS, Spiros. Reviewing Privacy in an Infomation Society. *Univsersity of Pennsylvania Law Review,* [s.l.], v. 135, n. 3, p. 707-746, 1987. Available at http://scholarship.law.upenn.edu/penn_law_review/vol135/iss3/3. Last viewed on Jan. 05.2018.

STOUGH, Roger; MCBRIDE, Dennis. Big Data and U.S. Public Policy. *Review of Policy Research,* [s.l.], v. 31, n. 4, p. 339-342, 2014.

TOMASEVICIUS FILHO, Eduardo. Em direção a um novo 1984? A tutela da vida privada entre a invasão de privacidade e a privacidade renunciada. *Revista da Faculdade de Direito da Universidade de São Paulo,* [s.l.], v. 109, p. 129-169, 2014.

UNICEF. The State of the world's children 2017. Children in a digital word. Available at https://www.unicef.org/publications/files/SOWC_2017_ENG_WEB.pdf. Last viewed on Dec. 12.2017.

WANG, Yinying. Big Opportunities and Big Concerns of Big Data in Education. *Techtrends,* [s.l.], v. 60, n. 4, p. 381-384, Abr. 27, 2016. Springer Nature. http://dx.doi.org/10.1007/s11528-016-0072-1.

WILLIAMSON, Andy. Big Data and the Implications for Government. *Legal Information Management,* [s.l.], v. 14, n. 04, p. 253-257, Dec. 2014. Cambridge University Press (CUP). http://dx.doi.org/10.1017/s1472669614000553.