Submission date: 15/01/2021
Resubmission date: 15/05/2021
Camera ready submission date: 10/07/2021

1st round notification: 22/04/2021
2nd round notification: 24/06/2021
Available online: 20/08/2021
Publication date: 20/08/2021

Section: extended version
Extended versions of selected papers from the IX Brazilian Workshop on Social Network Analysis and Mining

Social Network Analysis, Ethics and LGPD, considerations in research

Luiz Paulo Carvalho[1], Jonice Oliveira[1], Flávia Maria Santoro[2], Claudia Cappelli[1]

[1]Graduate Program in Informatic – Universidade Federal do Rio de Janeiro (UFRJ)
Rio de Janeiro, RJ – Brasil

[2]Computer Science Department – Universidade do Estado do Rio de Janeiro (UERJ)
Rio de Janeiro, RJ – Brasil

luiz.paulo.carvalho@ppgi.ufrj.br, jonice@dcc.ufrj.br,
flavia@ime.uerj.br, claudia.cappelli@gmail.com

Abstract. *Data protection and data-driven solutions are two progressing areas permeating Brazilian society. From an analytical qualitative interpretative research approach, this work presents an interdisciplinary study related to Ethics, from the ethics in computing perspective; the LGPD, from the Law studies perspective; and the Social Network Analysis in Brazil, from the Informatics perspective. This research area utilizes personal data extensively for knowledge construction, with semantic contributions, analyzing the reality; or pragmatic, building artifacts. Challenges and inseparable issues are observed, exposed, and debated in this work. As result and contribution, we present considerations combining the three topics, personal data in the research field of Social Network Analysis in Brazil respecting the LGPD and ethics precepts.*

Keywords. *LGPD, Social Network Analysis, Ethics in Computing, Data Protection, Personal Data-based Research*

Resumo. *A proteção de dados e as soluções orientadas a dados são duas áreas em progresso que permeiam a sociedade brasileira. Através de uma abordagem de pesquisa qualitativa interpretativa analítica, este trabalho apresenta uma abordagem teórica interdisciplinar relacionada à Ética, na perspectiva de Ética na Computação; à LGPD, na perspectiva dos estudos de Direito; e na Análise de Redes Sociais no Brasil, na perspectiva da Informática. Esta área de pesquisa utiliza dados pessoais extensivamente para construção do conhecimento, com contribuições semânticas, analisando a realidade; pragmáticos, na construção de artefatos. Desafios e questões inseparáveis são observados, expostos e debatidos neste trabalho. Como resultado e contribuição, apresentamos considerações combinando os três tópicos, dados pessoais no campo da pesquisa em Análise de Redes Sociais no Brasil, respeitando a LGPD e preceitos éticos.*

Palavras-Chave. *LGPD, Análise de Redes Sociais, Ética em Computação, Proteção de Dados, Pesquisa Baseada em Dados Pessoais.*

Cite as: Carvalho, L. P., Oliveira, J., Santoro, F. M. & Cappelli, C. (2021). Social Network Analysis, Ethics and LGPD, considerations in research. iSys: Revista Brasileira de Sistemas de Informação (Brazilian Journal of Information Systems), 14(2), 28-52. DOI: 10.5753/isys.2021.1235

1. Introduction

"The most valuable resource in the world is no longer oil, it is data" [The Economist 2017] announce the famous expression associated with Clive Humby, 2006, *"Data is the new oil"*. It foreshadows the influences of the theme in several sectors like academic and legal. Moreover, we observe the increase in computational capacity, such as data storage and processing, allowing procedures that were previously impossible or too much costly, such as multidimensional calculations based on attributes extracted from Online Social Networks (OSN) [Sumpter 2018]. However, from another side, there are cases, as the Cambridge Analytica [Isaak and Hanna 2018], which exposed an illegal and unethical use of data of the Online Social Network (OSN) Facebook users.

In the academic area, we can mention the Human-Data Interaction [Mortier et al. 2015], Data Ecosystems [Oliveira and Lóscio 2018], Data-driven Society [Pentland 2013], Data Economics [Börner et al. 2018], Data Science and Business Intelligence [Larson and Chang 2016], among others. Ethical issues emerged in this context, enabled by achievable high performance and distribution of computational technologies [Moor 2005], and due to discredit on data manipulation and its impacts on a society ignorant or naive of how its data is used. Governments have strengthened their data legislation, bringing those problems to light and taking them to a new level of importance.

The General Data Protection Regulation (GDPR) [EU 2018] has come into force throughout the European Union (EU) since 2018. GDPR has a transnational influence, involving EU countries and partners, such as Brazil. It determines that only countries with legislation as strict, or even more, can handle the personal data of EU citizens, even those with dual citizenship. Brazilian OSN, or other computer systems-based companies, could only handle data from European users if their specifications and requirements were compatible with GDPR. It includes the Social Network Analysis (SNA) research area in Brazil containing data from EU citizens.

To comply with GDPR directives and establish legal sovereignty over the data of its citizens [Bioni 2019], the Brazilian government approved the General Data Protection Law (*Lei Geral de Proteção de Dados Pessoais* - LGPD) [Brasil 2018]. The law came into force on September 18, 2020 and influences all data processing of Brazilian natural person [1], in Brazilian territory or abroad, by all its partner countries, similar to the European directive. The administrative sanctions, or punishments, are expected to take effect in August 2021. The relevance of the theme is observed from a Constitutional Amendment Proposal (*Projeto de Emenda Constitucional* - PEC) to add data protection to the Federal Constitution, setting the private competence of the Union to legislate on the matter [Brasil 2019b].

According to [Hijmans and Raab 2018], besides GDPR, we have been noticing a spread of ethical discourse about data protection, including the creation of ethical codes of practice, ethics advisory processes, and groups. While the laws define what must, can, or cannot be done, ethics is related to what is considered good and bad behavior.

[1] In the Brazilian context, and by LGPD terms, individuals are known as a natural person (*pessoa natural*); while entities formed by individuals and recognized by the State are known as a legal person (*pessoa jurídica*).

Data protection laws are based on ethical notions that reinforce the fundamental rights of privacy and data protection. GDPR embeds ethical principles, and ethical analysis is part of this law application since it contains several components that may require an ethical judgment and cannot be applied exclusively by technical approaches.

SNA area is affected by the panorama of data relevance in society, the LGPD scenario, and ethical dilemmas. Research involving a natural person data in the SNA area, whether obtained from OSN or not, will be under the auspices of the law, addressed by us in this article. OSN composed, for example, of animals, companies, or abstract elements are out of LGPD's scope.

According to the Digital 2020 Reports [2], Brazilians spend an average of 3h34m online on OSNs per day; Facebook Messenger and Facebook are the two most downloaded apps by Brazilians in 2019; Brazilian are third place in time usage of mobile Internet, 4h41m daily; 79% of Brazilians are concerned with how companies use their personal data; 85% of Brazilians are concerned about what is fake or real on the Internet; Brazilian number of OSN users increased by 11 million (+8.2%) between 2019 and 2020; One hundred and twenty million Brazilians are reachable by advertising on Facebook.

These data point to the significant influence of the Internet and of OSN in Brazil, reflected in the importance and intention of conducting academic research using these channels. The LGPD text is not simple or easy to assimilate, it was not written for researchers or computer scientists. In this paper we will contextualize the law with common practices in SNA, folloring the principles of a practical empirical proposal. We seek to answer: how is the interdisciplinary crossing and intersection of LGPD and the epistemological praxis of SNA research? What recommendations and guidelines can be drawn from this perspective?

The methodological approach is theoretical-analytical, detailing minutely the LGPD and relating to the State of the Art that SNA practices [Can et al. 2014, Yang et al. 2017]. Our objective in this paper is to cross the Legal domain and the Computer domain, correlating the praxis in SNA, with a potential process of personal data, and LGPD conceptual or technical details, developed with Legal hermeneutic and complex specificities distant to Computing discourse.

And what is the relationship between LGPD and the daily life of the Brazilian researcher? Research not compliant with legal proceedings is at risk of deindexation, exclusion, or administrative sanctions involving the researcher or the research institution, in extreme cases. Initially, it is already expected that there will be a certain lack of preparation on certain organizations [Bioni 2019], including research institutions. A period of maturity and transition towards legal compliance is also expected, we intend to prepare and anticipate interested stakeholders.

The work is structured as follows: Section 2 summarizes the theoretical foundations; Section 3 exposes the research paradigm and methodology; Section 4 presents the contribution; in Section 5 we forward the discussion; and Section 6 concludes this paper.

[2]https://datareportal.com/reports/digital-2020-brazil. Available in 01/01/2021

2. Theoretical Foundations

2.1. Brazilian Data Protection General Law (LGPD)

Law 13.709, approved on August 14, 2018, also known as LGPD [3] aims to protect the fundamental rights of freedom and privacy and the free development of the personality of the individual [Brasil 2018], who has his data in OSN, whether he is a user or not, e.g., published photos exposing users not registered in the OSN and yet processed, even if those users do not know or consent. The posterior Law 13.853, approved on the July 7, 2019, changed and added elements to the LGPD [Brasil 2019a].

The data protection discipline presented at the LGPD is grounded on the following basis [Brasil 2018]:

I – respect for privacy;
II – informative self-determination;
III – freedom of expression, information, communication and opinion;
IV – inviolability of intimacy, honor and reputation;
V – economic and technological development and innovation;
VI – free initiative, free competition and consumer protection; and
VII – human rights, free development of personality, dignity and exercise of citizenship by the individuals.

The principles that guide the data processing are listed in Art. 6. Preceded, above all as determined in the *caput*, by good faith. Table 1 presents the principles, widely discussed in this work, tangent to the SNA.

Data processing [4] is defined in an exemplary, non-exhaustive or limiting list of actions, direct or indirect, involving data. The organization where the research is carried out acts as **controller** [5], while the researcher processing the personal data may or may not play the role of **processor** [6]. They are organizational processing agents, whether public, private, third sector, among others. Controller and processor can be external actors to the organization, outsourced.

Studies relate the GDPR and European scientific epistemology [Chassang 2017]. Despite the laws' similarities, they are not identical and have different textual constructs for similar concepts. For example, GDPR does not explicitly point out the principle of non-discrimination and its definition for data processing and data holder are different from LGPD.

Regarding the LGPD and GDPR differences, we also point out the maturity of time in effect, the GDPR has been in force since 2018. GDPR has dedicated elements to

[3] We do not intend to scrutinize the LGPD or its terms. The definitions of the law terms will be presented in the footnotes, translated from Brazilian Portuguese to English by LGPD Brasil in cutt.ly/Zhd8pAI

[4] "processing: any operation carried out with personal data, such as those that refer to the collection, production, receipt, classification, use, access, reproduction, transmission, distribution, processing, filing, storage, elimination, information evaluation or control, modification, communication, transfer, diffusion or extraction;" LGPD, Art. 5, X

[5] "natural person or legal entity, governed by public or private law, in charge of making decisions about the processing of personal data;" LGPD, Art. 5, VI

[6] "natural person or legal entity, governed by public or private law, which process personal data in the name of the controller;" LGPD, Art. 5, VII

Table 1. Data processing LGPD principles

#	Principle and definition
I	**Purpose**: processing for legitimate, specific and explicit purposes informed to the data subject, without any possibility of subsequent processing inconsistently with these purposes;
II	**Adequacy**: compatibility of the processing with the purposes informed to the data subject, in accordance with the context of the processing;
III	**Need**: limitation of the processing to the minimum processing required for achievement of its purposes, encompassing pertinent, proportional and non-excessive data in relation to the purposes of the data processing;
IV	**Free access**: guarantee, to the data subjects, of facilitated and free consultation on the form and duration of the processing, as well as on all their personal data;
V	**Quality of data**: guarantee, to the data subjects, of accuracy, clarity, relevance and update of the data, according to the need and for compliance with the purpose of the processing thereof;
VI	**Transparency**: guarantee, to the data subjects, of clear, accurate and easily accessible information on the processing and the respective processing agents, subject to business and industrial secrets;
VII	**Security**: use of technical and administrative measures able to protect the personal data from unauthorized accesses and from accidental or unlawful situations of destruction, loss, alteration, communication or diffusion;
VIII	**Prevention**: adoption of measures to prevent the occurrence of damage in view of the processing of personal data;
IX	**Non-discrimination**: impossibility of processing data for discriminatory, unlawful or abusive purposes;
X	**Liability and accounting**: proof, by the agent, of adoption of effective measures able to prove observance of and compliance with the personal data protection rules, and also with the effectiveness of these measures.

punctually deal with academic research and data processing involving this practice, Art. 9, Art. 89, Recital 159 [EU 2018]. Some concepts and constructs from the LGPD were imported from the GDPR, equal or similar, and despite the differences, the tendency is that resolutions for Brazilian cases involving this theme follow similar and transferable European interpretations [Moribe et al. 2019].

Comparisons between LGPD and GDPR are already present in reports [Moribe et al. 2019] [Yun et al. 2020] but this topic is out of the scope of this paper.

2.2. Ethics, Research, and Personal Data

Extensive literature relates Ethics, Research and Science [Reijers et al. 2018], some contributions incorporated in this work. However, there is still no work that associates LGPD and SNA, considering the context and the Brazilian reality. Privacy, on the other hand, is not an unprecedented issue in SNA [Zheleva and Getoor 2011], despite this and until the LGPD there was not even such a comprehensive and specific general law specifically addressed to personal data, sensitive or not, in Brazil. Concerning general research, the Brazilian Association of Research Companies (*Associação Brasileira de Empresas de Pesquisa* - ABEP) has a simple and objective guide facilitating the association between LGPD and Research in a broad spectrum [ABEP 2017], but lacking an SNA approach.

Founded in 1947, the Association for Computing Machinery (ACM) is one of the organized and closed communities dedicated to computing research and practice, and its

ethical implications [7]. The first ethical guideline dates from 1966. In the second version, from 1972, it says: "EC5.2. An ACM member, whenever dealing with data concerning individuals, shall always consider the principle of the individual's privacy and seek the following: To minimize the data collected. To limit authorized access to the data. To provide proper security for the To determine the required retention period of the data. To ensure proper disposal of the data". In this excerpt we can clearly see an ethical concern with the processing of personal data and privacy, almost fifty years ago.

In Brazil, the entity equivalent to the ACM, in the USA, is the Brazilian Computer Society (*Sociedade Brasileira de Computação* – SBC), founded in 1978. SBC published its first Code of Ethics in 2013 [8] (forty-seven years after ACM), with writing resembling the 1966 and 1972 versions of ACM, and outdated compared to the 1992 version. The fourth and last version of the ACM Code of Ethics dates back to 2018, importing guidelines associated with contemporary dilemmas. Related to the interaction between non-governmental societies and ethical aspects associated with computing, there is a gap and immaturity of the ethical debate in the Brazilian scenario.

Concerning ethics related to GPDR, [Fabiano 2020] states that there is also the soft law that consists of opinions issued by Data Protection Supervisory Authorities and the European Data Protection Board, which provide clarification to support the interpretation of the data protection law. The author argues that it should be possible to apply the GDPR principles of thinking ethical: *"it is a matter of approach even without any norm"* [Fabiano 2020]. In this sense, he suggests that every single subject involved in the processing of personal data must be conscious of the high value belonging to a natural person such as human dignity; technical and organizational measures should guarantee "by default" the processing of personal information necessary for each specific purpose and each natural person should know her rights laid down by the data protection law.

When considering research involving SNA conducted in an organization configured as a research institute, and using personal data, there is a potential need for the involvement of a Research Ethics Board, Research Ethics Committee (*Comitê de Ética em Pesquisa* – CEP), or a similar entity. If required by the university, research involving human beings needs to be evaluated and approved by a CEP. In Brazil, these committees are more frequently found in the Health knowledge areas, with some exceptions in the Computer Science [Amorim et al. 2019].

2.3. Social Network Analysis

One of the essential practices in SNA involves the collection of personal data on OSN. In this context, the collection can occur involving few individuals, through questionnaires; or through automated extraction involving many individuals, through dedicated Application Programming Interfaces (API). OSN Twitter is widely used for SNA research through its API, as we can see in the annals of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM) and the Brazilian Workshop on Social Network Analysis and Mining (BraSNAM).

[7] https://ethics.acm.org/code-of-ethics/previous-versions/. Available in 01/01/2020.

[8] https://www.sbc.org.br/institucional-3/codigo-de-etica. Available in 01/01/2020 (in Brazilian Portuguese).

In this work, we use two compendiums as a conceptual foundation and as a basis for research practices in SNA, [Can et al. 2014] and [Yang et al. 2017]. These two works provide the computing technical framework.

Briefly pointed out the theoretical bases that guide this work, we proceed to the methodological approach and research method.

3. Methodological approach

In this paper, we discuss the intertwining between LGPD and research practices in the Social Network Analysis area. The LGPD is primarily directed to the Brazilian reality, we focus on this geolocation, considering it both as a public policy of data protection and privacy as a legal norm linked to the technological domain. We combine the discussion on Technology, Law, and Public Policies [Bucci 2008] and possible implications of LGPD in the Brazilian scientific, academic, and research scene, through a qualitative, interpretative and analytical interdisciplinary approach, considering the episteme of the domains of SNA, Law Studies, and Ethics. The normative world and the technological world are taken as unison elements.

We seek to fulfill the social function of empirical scientific research by adopting strengthening mechanisms that facilitate interested parties to adopt perspectives of structural change in society [Reis 2015], that is, to advance the debate and the practice of data protection, through the principles and bases of LGPD in the SNA domain. We seek to transcend the role of research subject observers, committing ourselves, with academic rigor and quality, to improve the reality of Brazilian scientific epistemology [Reis 2015]. In addition, we have extended the debate to the domain of Ethics, confident that the topic of data protection and privacy is not limited to the LGPD.

Indirectly, as analytical empirical research with a practical bias, we assist in the law dissemination, adoption, and compliance, guaranteeing its legitimacy of application. Voluntary acceptance is associated with values and virtues, not just use [Tyler 2017]. This perception is encompassed by the research method, where we seek to condense and summarize only the respective content. As qualitative research, we value depth grounded in practice and their respective values and virtues, without philosophical or abstract detours except for ethical considerations.

We use a methodological proposal similar to [Mello and Araújo 2015] when analyzing the Law of Access to Information through an interdisciplinary approach. There are two research objects in different knowledge domains, the LGPD and the practices in SNA within the scope of academic-scientific praxis. We associate direct information, as well as in the letter of the law; or indirect information, such as specialized literature on the topic; to point out analytically proposals for concrete, complete and objective recommendations on how to think-do research on SNA under the auspices of the LGPD. The contribution is twofold, both to the domain of computing for the elaboration of research in compliance with the law and to the domain of Law Studies to observe the panorama on how these two objects are related.

This research falls into the category of prescriptive analytics, where we supplant and extend the analysis to the point of suggesting future solutions and forwarding concrete

analysis, based on formal and structured reasoning [Babbie 1998]. As an interdisciplinary approach, we reconcile the epistemology of Applied Social Sciences Applied to Computing Science. We seek literature on data protection and privacy, linked to the LGPD, and literature on the practice of research in Social Network Analysis, we analyze, evaluate and extract the knowledge to structure the reference base of this work. It is worth considering that the sorting of information was restricted to the objects of this research, for example, although the fine as an administrative sanction is a topic normally present in LGPD-related communications, as it is in no way associated with research practices in SNA, it is disregarded. Realistically, we consider it an extrapolation of scope to point out that the fine for breaching the law can reach the amount of R$ 50,000,000.00 since ethical epistemic responsibility precedes any fine in academic-scientific practice. It is noteworthy, in the context of this paper, that a research paper can be de-indexed, discredited, or devalued, than to point out that there are potential millionaire fines involved.

The qualitative interpretative analytical approach is adequate to ensure depth, detail, and realistic and concrete positioning to the research. On the other hand, it is limited to the researchers' vision and understanding of the reality, supported by the quality of the references and analytical arguments.

4. LGPD and SNA, parsimoniously and ethically coexisting

In this Section, we detail and analyze the LGPD's most relevant points associated with SNA and ethical considerations. If applicable, we propose recommendations for research in this new scenario, considering the law in effect since 2020. The prescriptive analytical contribution in this work, recommendations, and notes between the LGPD and the epistemological praxis of SNA research are stitched together.

4.1. Principles and SNA research

First, we associate each principle presented in Table 1 with the common practices of SNA research [Can et al. 2014, Yang et al. 2017].

For the **principle of purpose**, the legislation will require a maturity in the convergence between the proposal and its research objective, its questions or hypotheses, and the data collected. How each of these connects and what is the purpose for processing each one. In qualitative or mixed research, this information is primarily important. In quantitative research, without individual interaction or objective permission, the recommendation is that data not related to the purpose of the research, related to principles III and VII, should not be kept. For example, there was an data extraction from Facebook, returning many dimensions; is the number of reactions to a particular publication compatible with the research objective, or is this incompatible or useless data? In the case of useless data, it should be discarded.

By the **principle of adequacy**, other aspects then purpose must be considered. Time is one of them, personal data is not owned by the researcher so that it can be kept "forever". Security and access criteria must also be adequate, only the researchers initially determined, either in the research protocol or in the informed consent form, must have access to personal data; the available and reasonable security requirements must be

enabled. For example, the researcher notifies the individual that his laboratory will have access to his personal data, and another member of the laboratory shares all databases with third parties without analyzing the other access restrictions, it configures an adequacy violation, even if there is good faith. From this point on, the responsible researcher no longer has access to these same personal data that he has agreed to protect.

The **principle of necessity** is closely related to the purpose principle, it is also known as "data minimization". Just keep the minimum necessary, useful, and associated with the specific research. Too much data increases the responsibility of the researcher, or his research body, and makes control and security more difficult. For example, Harris [Harris 2015] shows how it is possible to extract ethnic information from individuals through their name, and ethnic data is classified by the LGPD as sensitive.

The **principle of free access** is trivial in qualitative research where the number of participants is small or medium, i.e., the participants are easily tracked and, therefore, their respective data easily found. In quantitative research, where clear and explicit identifications are not always present, this is still an open challenge. For example, in a database extracted from Twitter, a person aware of the research can request access to the personal data related to it, how to operationalize this request on an base with more than one million records, or involves more than one database? Since it may contain other people's personal data and the option "send the entire database" is not viable.

The **principle of quality of data** is imperative to the quality of the respective research. Research that uses erroneous, invented, or falsified data (as far as its essence is not based on this category of elements) is not scientifically adequate or valid.

The **principle of transparency** is crucial to the initial phases, mainly qualitative. The data collection must follow the quality criteria stipulated in this principle. After this stage, also the operationalization of the communication with the respective participants, they can ask to access (principle IV), modify (principle V) or delete the personal data associated with them. In this case, the best way is to maintain an open and accessible communication channel with the participants, also building transparent communication and interaction.

The **principle of security** is also simple. It is not the scope of this research to elaborate on this principle, it is the object of techniques in Information Security. The researcher must be aware of the mechanisms, in a simplified way, that will safeguard personal data. For example, if the data is stored in a password-protected Google spreadsheet, this must be reported.

The **principle of prevention** is already in line with the ethics in research established by CEP in Brazil since 2012 [Brasil 2012]. For example, in very sensitive cases, such as individuals in witness protection programs or HIV-positive people, the communication of additional prevention measures can make the person's perception of participation less dangerous or threaten for itself. The benefits of an ethical stance are twofold. The participants feel safe, legitimately involved with the research; and the researcher obtains reliable, spontaneous and genuine data.

The **principle of non-discrimination** is a differential between the LGPD and the

GDPR, and ensures that even in wide-ranging and publicized scientific research there is no damage to the personality, dignity, or humanity of the human person involved. For example, an ANS research to identify the sexual orientation of its participants, without consent or where the consent was poorly specified, can result in harm to the person involved, which is not the researcher's intention.

The **principle of liability and accounting** can be achieved if all the others are adequately achieved. What the researcher communicated and agreed with the participants must be respected through mechanisms and operational guarantees, otherwise, it will be subject to administrative sanctions.

4.2. Data Collection

Information entered by users in OSN is not public domain or public or open data, each platform defines ownership and rights in its terms, respecting (as expected) the specific laws where they operate [Donahue 2016].

Bond et al. [Bond et al. 2013] conducted a survey with forum users related to what they produce in virtual environments, on the perception of their products as public domain or not. There is no consensus, users are divided between those who are satisfied that their products can be used by anyone as they wish, while others prefer to have authorship and control of what they produce.

Regarding the LGPD, accompanying the GDPR effectiveness, if an individual realizes that a communication's object damages his privacy or conflict with articles 2 or 6, may require a response to be taken accordingly. Two LGPD mechanisms protect the research and the researcher: **consent** [9] and legitimate interest.

The researcher, when the research involves a small or reasonable number of participants, can use a free and informed consent form. This approach is more frequent in qualitative research [Recker 2013], which may involve **sensitive personal data** [10]. Sensitive personal data has a law section dedicated to it. Section II, and art. 13 deepens the use of these data in public health studies, which may or may not involve SNA. The term must follow all the principles listed in Table 1, enacting the purpose, processing time, storage location, among other information that must be disclosed, transparent. In the case of a relevant change in any of the research terms, new consent must be obtained, such as the transmission of personal data to another research group.

The legitimate interest [Mulholland 2020], art. 10, is an output for the treatment of a huge amount of data, as it often happens in SNA. *"The legitimate interest of the controller can only support the processing of personal data for legitimate purposes, considered based on specific situations, which include, but are not limited to: I - support and promotion of the controller's activities;"* [Brasil 2018]. Kotsios et al. [Kotsios et al. 2019] conducted a survey extracting Twitter data, with no success in obtaining the consent of

[9] "free, informed and unequivocal pronouncement by means of which the data subjects agree to the processing of their personal data for a specific purpose;" LGPD, Art. 5, XII

[10] "personal data on racial or ethnic origin, religious belief, public opinion, affiliation to union or religious, philosophical or political organization, data relating to the health or sex life, genetic or biometric data, whenever related to a natural person;" LGPD, Art. 5, II

each of the users holding the data they collected. Even though using an automated approach, Twitter stopped the issuance of direct messages requiring consent, alleging spamming practice. This case shows the current impossibility of obtaining widespread consent in research with an enormous volume of data from OSN.

If there is no free and informed consent and if the research data or information violates the grounds and legal principles, the stakeholders involved are liable to the administrative sanctions provided for by law, both processing agents. There will be intercession on the scientific communication, case by case. This same referral applies to surveys carried out by autonomous natural persons for exclusively private and non-economic purposes, where there is an infringement of the LGPD and wide public disclosure containing personal data, sensitive or not.

The LGPD is not applied for the processing of personal data made by a natural person for exclusively private and non-economic purposes, as determined by art. 4 [Brasil 2018]. There is an infraction, detailed in Section 4.3 when there is wide public disclosure containing personal data, sensitive or not, intentional or not. In this sense, we need to go beyond the LGPD, promoting education and ethical awareness regarding the processing of personal data by individual natural persons. Phenomena such as cyberstalking, dataveillance, social engineering, embarrassment, identity theft are fed and facilitated by personal data configured as public or opened by the Internet, harmful to society and especially to minorities, such as women [Carvalho et al. 2020].

For example, a Brazilian deputy conducted an SNA-like study and structured, based on personal data in OSN, a dossier on potential "anti-fascist" Brazilians, also addressed by him as "potential terrorists" [Fleck 2020]. As the LGPD determines, political positioning is considered sensitive data, to begin the analysis of this absurdity. These data were sent directly to law enforcement agencies and, indirectly, leaked on the Internet, with easily identifiable personal data. It was made by a natural person, with no financial purpose, for private interests, and, for some unknown reason, "leaked" to the Internet.

4.3. Protecting data primarily, people consequently

Even if free and informed consent and scientific communication make use of personal data, sensitive or not, potentially harmful to the data subject, a likely infraction occurs. One of the essential epistemological essences of the LGPD is the reinterpretation of the "protection" paradigm, where legislation protects data, individual or collective, aiming to achieve its objectives [Bioni 2019]. This interpretation is urgent when dealing with sensitive personal data, which can cause social damage to the data subject and violates the principle of non-discrimination.

Currently, the majority of the population is lay or naive regarding the processing of their personal data, and may not be aware of the magnitude of this practice. It can not discern the damage that will be caused to them through these complex or obscure computational solutions, i.e., the effective impacts or influences of that data processing from the moment that it consented to its treatment. This perception is supported by art. 42, II, § 2, allowing for the burden of proof inversion in favor of the individual when, by judgment, the allegation is credible, there is a failure to produce evidence [Brasil 2018],

and by the art. 6 principles, presented in Table 1.

Disinformation, such as fake news, and potentially controversial information using personal data, whether sensitive or not, is disseminated and spread quickly and uncontrollably in OSN [Avelar 2019], this behavior is intensified in digital non-native or less digitally literate Internet users [Guess et al. 2019]. An example is the pandemic COVID-19-related misinformation transmitted and shared through the OSN YouTube [Bhatta et al. 2020]. Published research data can be legally reinterpreted and appropriated by media and disseminated to unexpected proportions, putting those involved at risk, consenting or not.

Some sensitive information, data with semantic injection, can be discovered when analyzing proximity and neighbors in networks. With phenomena such as the Bubble Effect or Echo Chamber, certain networks can make individuals identifiable as well as expose sensitive data about them. This scenario is easy to see in genetic analyzes, e.g., involving DNA; when a person exposes genetic information, he simultaneously delivers genetic information from his family. And these family members may not have consented. Another example, consider an environment or event categorized as LGBTQIA+, building a network associating an individual with this environment or event can expose their sexual orientation. These examples point out how this is a collective issue, not an individual one; and how they are related primarily to data, secondarily to individuals [Véliz 2020]. Scientific research, even under the cover and supposed immanence, can end up exposing sensitive data. The researcher must have maturity on the personal data externalized through the research, aware of the possible subsequent ethical implications if any.

4.4. Academic, scientific or study purposes

The law touches on academic, scientific and studies purposes in three situations: (i) Academic purposes (art. 4, II, b): "Academic purposes, in which case articles 7 and 11 of this Law shall apply;" [Brasil 2018]; (ii) Conducting studies (art. 7, IV and art. 11, II, c). "For the conduction of studies by **research bodies** [11], guaranteeing, whenever possible, the anonymization of personal data" [Brasil 2018], art. 7 deals with personal data and art. 11 deals with sensitive personal data; (iii) Research body (art. 5, XVIII, art. 7, IV and art. 11, II, c). Research, involving academic or study purposes, is limited to these areas in the LGPD. In addition to these three, art. 4, I: "This Law does not apply to the processing of personal data carried out by a natural person for exclusively private and non-economic purposes" [Brasil 2018].

[Belmudes 2020] analyze the relationship between LGPD and OSN, including the scientific communication scope. The European Court of Justice (ECJ) reinforced and ratified its position regarding the use of personal data, sensitive or not, coming from social networks, "even if a personal data is used in a particular and not economical way if its processing gives visibility to a number indefinite number of people, the exception does not apply" [Belmudes 2020]. Considering that a research body like the one considered

[11] "body or entity of the direct or indirect public administration or not-for-profit legal entity governed by private law organized under the Brazilian laws, with its principal place of business and jurisdiction in Brazil, which includes in its institutional mission or its corporate purpose or purpose established in the By-Laws basic or applied historic, scientific, technological or statistical research;" LGPD, art. 5, XVIII

in art. 5, XVIII, has non-profit purposes, as far as the economic aspect is concerned, a scientific communication made available and broadly disseminated to society is under LGPD scrutiny.

Therefore, if a researcher develops and assembles widely available tacit scientific communication that violates the LGPD grounds, art. 2, the disagreement with the principles will be observed in Table 1. The data subject, in this case, may request the elimination of personal data, art. 18, which identifies or lead to identification, which may result in the exclusion or deindexation of the research object as a whole.

For example, art. 4 can lead to the interpretation: as the law does not affect academic research, the personal data collected, sensitive or not, can be transferred between parties freely, and are exempt from non-functional security requirements or release the exposure of the holders. This understanding is absurd, stating that even purely academic purposes must be guided by the LGPD.

Art. 4 and its determination to "shall not apply" for academic purposes may culminate in a backdoor for unethical use by organizations. It is determined that the processing of personal data for academic purposes respects Art. 7, generic personal data, and Art. 11, of sensitive personal data. In Art. 7, IV, and Art. 11, II, c, the wording provides an exception, anonymization is not mandatory for all cases and its guarantee is subjectively delegated to the "whenever possible" criterion [12]. What is understood as "possible", then, must be negotiated between processor and controller, between researcher or research group and entity officially defined as controller. We stress this specific point because of the possible interpretive openness, and to reinforce the ethical respect for data protection by researchers and their respective research.

Personal data are owned by individuals to whom there is identified or identifiable traceability, as determined by art. 17 of the LGPD [13]. For example, a controversial and structurally unique publication on an OSN, a video on YouTube, or a photo on Instagram. This article reinforces the association between the individual, his personal data, and his digital production.

Then we return to the quantitative, qualitative, or mixed nature of SNA research. Even identifying your research as quantitative, using statistical, positivistic, structuralist, objective, and deterministic approaches, it does not automatically and sufficiently result in its positioning as research that does not involve identified or identifiable human beings through instrumentalized personal data [Zheleva and Getoor 2011]. For traditional qualitative research and identified subjects, CEP evaluation is mandatory, regardless of LGPD [Amorim et al. 2019].

A strong criticism to the operation of CEPs is their immanent distancing from the specific scientific epistemologies of the so-called technological courses, where we can perceive a Human-Computer Interaction aspect. Will a CEP comprised only of health

[12] "the conduction of studies by research bodies, guaranteeing, whenever possible, anonymization of the sensitive personal data;" LGPD, art. 7, IV; art. 11, II, c

[13] "All natural people are ensured the ownership of their personal data and the guarantee of the fundamental rights to freedom, intimacy and privacy, pursuant to the provisions of this Law" LGPD, art. 17

specialists be able to analyze and evaluate technological specifications and techniques for SNA research [Amorim et al. 2019]? Through this consideration, we propose a specific CEP for research involving human beings with a predominant technological position, or in the last case that at least one component of the CEP has a high level of specialization in computational technologies involving human beings.

4.5. Sensitive personal data, health-related or not

The LGPD is austere and objective regarding the security of information involving sensitive personal data and, emphatically, sensitive personal data related to health, art. 11, II, g, § 4 and § 5; art. 13. An example in the case of health-related data are studies of Epidemiology and ARS [Stattner and Vidot 2011]. Consent to use this category of data must be specified in a specific and detailed manner, not obscured in a generalized way.

In works involving this category of data, care and precaution must be primary research requirements, so as not to allow reverse tracking of data to respective individuals, identifying them, offering risk to them and the research.

In Russia, a student was expelled from the educational institution for having his personal data from his OSN analyzed, categorized as "gay" by the director board [The Moscow Times 2019]. It configures moral harassment and unscrupulous use of data. OSN Russian users are being pursued and watched by government officials. SNA approaches are used for these purposes, not only isolated individuals are targeted, but their networks of contacts [Gabdulhakov 2020]. These two cases superficially demonstrate the damaging potential that an SNA study can generate, even in a country where data protection legislation exists, such as Russia [Russia 2006].

In scenarios where there is political persecution, discriminatory extremism, or data surveillance, such as Brazil during the Bolsonaro government [Kemeny 2020], SNA can be a powerful and ethically flexible mechanism, we can quote:

I Use of quantitative and, mainly, qualitative approaches to probe and control data of individuals in OSN. A clear example is a dossier, entitled "map of influencers" [Valente 2020], ordered from an agency in the form of research or study, SNA-based research. Eighty-one people were registered and profiled in this dossier, without any consent, exposing data such as their phone numbers and e-mail addresses [Valente 2020]. Actions were proposed, such as "preventive monitoring" to those considered "detractors", and "proposing a partnership" to "allies". The list includes several professors and researchers.

II Hypothetically, when one of these professors or researchers exposes scientific research involving ANS, quantitative or qualitative, in their broad communications, if they are "under preventive monitoring" there is no guarantee or democratic security that the censors will not delve into the research exposed, or invade the privacy of the researcher in question. This is just a possible real case based on the concrete situation of the Brazilian communicational environment. The research is under the legal scrutiny of the LGPD, and due to these censorship initiatives, the researcher himself and the knowledge generated by the research may become the target of an investigation by third parties, with malicious intentions.

4.6. Anonymization, pseudonymization and direct or indirect identification

LGPD deals with **anonymized data** [14], **anonymization** [15] and **pseudonymization** [16]. Personal data that has not been processed using an anonymization or pseudonymization approach effortlessly associates an individual with respective data and vice-versa, identifying it, directly or indirectly.

Direct association happens when the specific personal data is the only data needed to track and identify the individual, such as Individual Registration (*Registro Geral* - RG) or Natural Persons Register (*Cadastro de Pessoa Física* - CPF). Indirect association happens when the data alone is insufficient for the association, but a combination of data or semantic injections does the job. Indirect identification is not trivial and requires specific procedures, for example, to infer the sexual orientation of users in an OSN based on their interactions and production [Jernigan and Mistree 2009].

Pseudonymization is mentioned only in art. 13, specifically for public health studies, and health-related research communication. Then there is the recommendation that these studies follow the ethical standards related to studies and research, without specifying which ones. Concepts related to Ethics have only one occurrence in the text of the law, at this point. In addition, the interpretation of "good faith" can be found in the *caput* of art. 6, distancing itself from the universe of classical ethics and approaching the universe of moral oriented by the instance, or set of them.

Art. 12 determines: "Anonymized data shall not be deemed personal data for purposes of this Law, except when the anonymization process to which they have been submitted is reversed, using solely the appropriate means, or whenever it can be reversed with reasonable efforts." [Brasil 2018]. Therefore, the recommendation is that anonymization should be the rule and association, direct or indirect, the exception. Reasonable anonymization guarantees the non-incidence of LGPD.

The LGPD incorporates the concept of "reasonable technical means". Considering the impossibility of building an infallible or unbreakable solution and that anonymization is a complex activity. A small Faculty or research institute that researches in SNA will not have the same resources available for database anonymization compared to a large University.

Anonymization is also a recommended way out when reaching the data processing period agreed by the individual. If the individual cannot be identified or identifiable, there are no law violations. The LGPD determines that after all the purposes involving the treatment have been carried out by the research or study in question, the data should be deleted. Anonymization is a viable option, as it is no longer considered "personal data".

Qualitative research that depends on the identification of natural persons must take

[14] "data relating to a data subject who cannot be identified, considering the use of reasonable technical means available at the time of the processed thereof;" LGPD, art. 5, III

[15] "use of reasonable technical means available at the time of processing, by means of which the data loses the possibility of a direct or indirect association to a natural person;" LGPD, art. 5, XI

[16] "[...] the pseudonymization is the processing by means of which a data loses the possibility of direct or indirect association to a natural person, except for the use of additional information separately kept by the controller in a controlled and safe environment." LGPD, art. 13, § 4

extra care since the (or one of) the object of the research is an individual and their personal data are protected by the LGPD. A possible alternative is to disclose the work to the data subjects before submitting it to actual externalization, demonstrating that the individual owns its personal data and ensuing inferences.

A mathematical alternative is to use differential privacy algorithms to protect data, being a complex and advanced option, not suitable for simple or reasonable cases. Vadhan [Vadhan 2017] discusses this topic from an introductory level. Given an output, the observer is unable to identify the computed information specific to a given individual through mathematical procedures, said to be differentially private. In the US, large agencies release demographic information using differential privacy approaches.

4.7. Privacy, public servants, and personal data

There is still no consensus or definitive note regarding the so-called "private" communications in OSN issued by civil servants in official profiles, personal or private. Certain surveys that use the SNA episteme use profiles of politicians or civil servants, and their contents, to answer questions about reality. To illustrate, we propose some questions: the mayor of a certain city has a Twitter account, which is an OSN, configured as closed and "private"; are the messages produced by him during his office hours considered public because he "is" mayor in that time interval, or are they considered public full time because he "is" full-time mayor? Considering that the idea of the LGPD is precisely to bring light and to scrutinize the specifics of data protection, with this digital privacy, should not the law favor him with isonomy? His account configuration is considered closed and "private", does this not constitute legitimate good faith in restricting the wide publication of its own data through the available technical means?

The reasons for the activities carried out by public servants are different, especially about the limits of privacy rights, which make them different from not public servants [Ruviaro and Nedel 2017]. Public agencies have already developed manuals relating to the use of civil servants and social media, including that of the Secretariat for Social Communication to guide members of the federal executive branch to act on social media [Brasil 2014]; and the manual of good practices in the city of Rio de Janeiro [Rio de Janeiro 2018]; some others have been found, but are derived from these. These can be transferred or adapted to other contexts and spheres, to instruct public servants in basic notions of data protection and digital ethics in OSN.

[Ruviaro and Nedel 2017] point to an illegal case involving a public servant from 2016. A judge published images carrying out political party activities, campaigning for the determined political party, on Facebook. Data collected in cases like this are fortuitous for qualitative or mixed studies of egocentric social networks [Perry et al. 2018], a strand of SNA, specially in social sciences.

4.8. Profiling and specificity of terms

One of the LGPD concerns, as the GDPR [Kotsios et al. 2019], is the construction of deterministic personal data-related profiles, especially sensitive ones. This practice is known as profiling. The research develops profiles or models individuals. This is still an open question, not addressed by the law because research can build profiles that will bring

negative consequences to the holders themselves. What are the limits of this research? Ethical standards? Is possibly harming one or a few people a necessary and sufficient reason to slow down the research process? What are the limits of this profiling?

Specificity of terms will be two-pointed, defective consent and generalization of terms. Defective consent occurs "if the person, that is, the declarant had real knowledge of the situation, he would not have manifested his will in the way it was declared." [Lobo 2015], if the individual knew that his sexual orientation would be used in an SNA research, would s/he publish this data on Facebook? If an employee knew that the personality test s/he carried out at the company, through a private organization of organizational research, could harm the job integrity, could s/he deny the participation? Without retaliation? If a Twitter user knew that their posts are used for behavioral analysis SNA research, would he still post the same things? Or, simply, would you still publish? The terms of consent and justifications of legitimate interest must be understandable, clear, and simplified.

The generalization of terms refers to art. 8, § 4, "Consent must refer to specific purposes, and generic authorizations for the processing of personal data will be void." [Brasil 2018]. The data processing cannot be based on generic, eternal, or misleading purposes. In this sense, it is a challenge to expose the terms and requirements of the research in a transparent, informative manner, respecting the understanding of non-functional requirements [Carvalho et al. 2019]. The researcher cannot condense all consent to "research" or "study". Conducting scientific research must meet methodological rigor and specific protocols [Recker 2013]. The researcher must be aware of the data involved, the purposes, treatment period, and storage environment, especially because he will not be processing something that is his or her property, rather something that has been designed limited processing right according to the research criteria and protocol.

Covered all the points interpreted analytically as pertinent of prescription in qualitative synthesis, we proceed to the discussion.

5. Discussion

Despite the apparent rigor that seems to affect the data treatment exposed by this work, this topic is still immature and incipient in Brazil, also denoting the innovative character of this contribution. Many issues that have not yet been addressed in the LGPD, are open issues, or open room for opportunistic or obscure interpretations, especially in the academic, study, and research fields. In this sense, the recommendation is not to neglect the scenario and position yourself by pretending not to understand what is happening to take advantage or to guarantee your own good; rather, taking the reins of this academic agenda and building solutions that guide the use of personal data in SNA in Brazil, maturely and ethically.

Taking the reins of the debate and the construction of the ethical narrative related to computational processed data, in this case, more than a symbolic connotative recommendation, it is a concrete call to the ethical dialogic related to data processing not only with the Brazilian technological community but also with reports and opinions external to it. As explained by Moor [Moor 2005], the absence of requirements or ethical scrutiny in

the development of new or maintenance of existing computer technologies will escalate to judicialization due to their influences, or it will occur quicker. The unscrupulous processing of data by some entities culminated, not exclusively, in the ideation, elaboration, and promulgation of the LGPD.

How did the education, at any academic level, raise awareness and ethically instruct those who integrate it that personal data should be, even if ideally, protected? Are courses on the various topics of Databases in Brazil considering the topic of protection of personal data in their technical presentations? A course or program where a graduate of formal computer education is not educated in the protection of personal data given the technological paradigm in the second decade of the 21st century, offers the minimum and prepares a professional to work in today's society? Here we do not plead for a "Computer Legislation" class, as is commonly found in Brazilian Computing curricula, because the ethical debate goes beyond the LGPD, with the LGPD being a response from the democratic system to a perceived threat [Moor 2005] . The ethical essence of data protection must be rooted in Brazilian Computing curricula and didactics, not segregated from it. It must be integrated and crossed in thinking-doing technological disciplines, not delegated to an isolated and disconnected one which will most likely be conducted by a specialist from an area outside Computing. Only then, more than personal data protection awareness in applications and computer systems development; research, whether quantitative, qualitative, or pragmatic, will be guided, in essence, by the same ethical principles.

The GDPR presents articles and referrals for academic purposes, which can be imported into our legal scenario. For example, in art. 17 and art. 89 [EU 2018] there is an exception related to the Right to Forgetfulness [Bioni 2019] and exclusion of personal data legally exploited in research, i.e., in the case of the legal treatment of personal data, the holder cannot request their exclusion, since this can de-characterize or ruin the research.

Many open questions can only be answered by the National Personal Data and Privacy Protection Council (*Autoridade Nacional de Proteção de Dados* - ANPD). The ANPD was formed very shortly after the law came into force, its board was approved on October 20, 2020. It was expected that it would be ready and operational in advance, to encourage the adaptation to the law by society, as indicated in its list of competencies, Art. 55-J. Cases omitted from the LGPD's wording or conflicts of interpretation will be decided by the ANPD, such as those involving academic purposes.

5.1. Open issues and research oppotunities

Some points are open questions and will be academic challenges regarding SNA research, this Section seeks to bring some of these and point out referrals.

Given the art. 7, § 4 [17]:, information configured as "public" or without any limitation of scope in OSN is data legally made manifestly public by the respective individual? Or, as Bioni [Bioni 2019] points out, should data processing respect the principle of the

[17] "The requirement of the consent set forth in the head provision of this article for the data manifestly made public by the data subject is waived, provided the rights of the data subject and the principles set forth in this Law are observed." LGPD, art. 7, § 4

purpose of the platform or system on which it is initially located?

How does LGPD deal with the processing of personal data already collected before it came into force? With terms of consent already concluded? In the EU, terms for data collection, such as consent, before the GDPR's term could only be considered valid if they conformed with the GDPR, otherwise, it was mandatory to elaborate and make available new terms, and to obtain new consent from all holders [EU 2018]. Will LGPD retro act to research already published? How should research bodies deal with personal data, sensitive or not, already stored in their repositories? That is, if this data is used for a new processing purpose, as there was no previous treatment, does this practice violate the LGPD?

How to differentiate data processing for research or studies in, or for, public or private organizations? For example, Uber's research sector is not considered a research body by LGPD, as defined by law. If Uber and a research body sign a cooperation agreement, will Uber be able to use the "non-applications" provided by LGPD to process personal data with the subterfuge that there is a scientific purpose? Respecting transparency and fairness requirements, how to sign these agreements without violating the LGPD?

How to instruct the respective data subjects to act if they identify illegal use of their personal data in an academic communication or study? If there is an organic relationship of data collection between the researcher, the idea is that at this moment the participant is informed of his rights by the LGPD, e.g., the research questionnaire already anticipates the terms of the law. In an inorganic relationship, such as automated data extraction, the researcher must persist in the legal determination that the personal data is owned by the individual and that this individual can claim the terms of the law from the researcher. It is crucial to remember that the absence of identified or identifiable personal data exempts research from involvement with the LGPD.

In the LGPD there is a specific section for children and adolescents. As was already understood by the CEPs [Brasil 2012], and reinforced by the LGPD, the processing of personal data of children and adolescents is associated with the consent of their legal guardians. In the case of qualitative research, the CEP guidelines are sufficient at first; in the case of surveys that collect data indiscriminately with a hefty volume and at least one child or adolescent may be in the records, as in Twitter extractions, this is still an open question. The relationship between effort and compensation is disproportionate if the solution is to scan the entire database and analyze entry by entry, or sets of them, on bases with hundreds of thousands or millions of entries.

6. Conclusion

In this work, we relate the scientific epistemology of ARS and ethics concerning the LGPD, in the Brazilian context where the legislation applies. The relevance of this contribution is prompted by the awareness of the use of personal data, sensitive or not, by SNA researchers interested in this category of entries in the development of their research. Failure to LGPD compliance leads to the potential possibility of research impairment, under the responsibility of the researcher and the institution in question, which is the data controller and the consenting party, legal or illegal, of personal data processing.

The recommendation to take the reins of the speech, ethically and responsibly, is, in addition, to create basic content that the ANPD, and other interested parties, use as a guide for their decision making and content formulation.

The construction of research or study epistemologically based on ethical non-functional requirements is a specific research area [Manzoor 2017], not trivial. Considering the complexity, or even impossibility, of reaching a level of data security where, for example, anonymization is irreversible, LGPD uses reasonableness of efforts and technologies to apply its terms. The legal, ethical, and good faith conduct is in the concrete application of the best available means of anonymization, as opposed to neglecting them under the thought of "this will be very easy to revert, never mind".

As a limitation, even though it is often cited in this work, we do not address the non-functional requirement of Security. This is an extensive topic and future work is needed to trace excerpts from the law with the Information Security episteme. [Sherwood et al. 2005] contextualizes the ITC Security domain with a Software Engineering matrix, the LGPD mainly encompasses the contextual and conceptual levels, that we deal with. Through the logical level and the respective levels below, the technical security measures superficially ordered in legislation, which are essential to the concrete practice of "protect data", are techno-materially implemented. We mean, when the LGPD announces "reasonable and available technical means", the technologically normative discourse ends at the conceptual level, and it will be up to the treatment agents to define specific technological instances considered "reasonable" and "available", in a reasoned and formalized manner.

The Brazilian government has prepared and published a guide called the "Security Framework Guide" [Brasil 2021] to assist in the implementation and dissemination of the data protection culture at a technical level, with an emphasis on LGPD. Even so, at the end of the Introduction it states:

> "In the same sense, it should be remembered that the adoption of this guide does not necessarily indicates complying with Brazilian legislation on security, privacy and protection of personal data, in particular Law No. 13,709, of August 14, 2018 - General Protection Law Personal Data (LGPD). However, this document can surely help the adopting institution to achieve the objectives set out in the related standards, by allowing the visualization of the maturity of its information security and data protection and by expanding the implementation of best practices on the topic." [free translation] [Brasil 2021]

Two threats to validity are observed. Internally, the approach used here allows us to discuss the topic in-depth and referential breadth, it does not exhaust the debate or covers all possible understandings, restricted to our interpretation. Due to the unprecedented character and very specific nature of the analysis, there is a significant limitation in the categorically close results through the literature review. Externally, this is qualitative research linked to its respective time and space of discourse, so any changes in the LGPD, or in the legal scenario of data protection and privacy, as well as in the epistemological praxis of research in SNA represent threats to certain prescriptive referrals exposed here. Even so, good practices and ethical debate are relevant and resistant to this threat.

As future work we point out the forwarding of the questions presented here; elaboration of more interdisciplinary works containing perceptions and theorizations between

the areas of SNA and Law, dialogically; construction of artifacts that assist researchers in complying with the LGPD in their research, as guides or recommendation systems; in-depth analysis of the treatment of personal data of children and adolescents, being a specific topic and highlighted even in the legislation.

7. Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

This is an extended version of the paper "Pesquisas em Análise de Redes Sociais e LGPD, análises e recomendações" published and awarded on the IX Brazilian Workshop on Social Network Analysis and Mining (BraSNAM).

References

ABEP (2017). GUIA DE PROTEÇÃO DE DADOS. LGPD para Profissionais de Pesquisa de Mercado, Opinião e Mídia. Available at: `cutt.ly/FtmXfA6`. Accessed: 05/05/2021 (in Brazilian Portuguese).

Amorim, P., Sacramento, C., Capra, E., Tavares, P., and Ferreira, S. B. L. (2019). Submeter ou não meu projeto de pesquisa em ihc ao comitê de Ética, eis a questão. *Proceedings of the XVIII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais*. (in Brazilian Portuguese).

Avelar, D. (2019). Whatsapp fake news during brazil election 'favoured bolsonaro'. Available at: `cutt.ly/DtmXD5D`. Accessed: 05/05/2021.

Babbie, E. R. (1998). *The practice of social research*. Wadsworth Publishing Company, Belmont, 8 edition.

Belmudes, G. (2020). Aplicação da lgpd nas redes sociais. Available at: `cutt.ly/EtmXJie`. Accessed: 05/05/2021 (in Brazilian Portuguese).

Bhatta, J., Sharma, S., Kandel, S., and Nepal, R. (2020). Infodemic monikers in social media during covid-19 pandemic. *Asia Pacific Journal of Health Management*, 15(4):95–97.

Bioni, B. (2019). *Proteção de Dados Pessoais. A Função e os Limites do Consentimento*. Forense, Rio de Janeiro, RJ, 2nd edition. (in Brazilian Portuguese).

Bond, C., Ahmed, O. H., Hind, M., Thomas, B., and Hewitt-Taylor, J. (2013). The conceptual and practical ethical dilemmas of using health discussion board posts as research data. *J Med Internet Res*, 15(6):e112.

Börner, K., Scrivner, O., Gallant, M., Ma, S., Liu, X., Chewning, K., Wu, L., and Evans, J. A. (2018). Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. *Proceedings of the National Academy of Sciences*, 115(50):12630–12637.

Brasil (2012). Ministério da saúde. RESOLUÇÃO CNS Nº 466, DE 12 DE DEZEMBRO DE 2012. Available at: `https://conselho.saude.gov.br/resolucoes/2012/Reso466.pdf`. Accessed: 05/05/2021 (in Brazilian Portuguese).

Brasil (2014). MANUAL DE ORIENTAÇÃO PARA ATUAÇÃO EM MÍDIAS SOCIAIS IDENTIDADE PADRÃO DE COMUNICAÇÃO DIGITAL DO PODER EXECUTIVO FEDERAL, v.2. Available at: http://www.secom.gov.br/pdfs-da-area-de-orientacoes-gerais/internet-e-redes-sociais/secommanualredessociaisout2012_pdf.pdf. Accessed: 05/05/2021 (in Brazilian Portuguese).

Brasil (2018). LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. Available at: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Accessed: 05/05/2021 (in Brazilian Portuguese).

Brasil (2019a). LEI Nº 13.853, DE 8 DE JULHO DE 2019. Available at: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Lei/L13853.htm. Accessed: 05/05/2021 (in Brazilian Portuguese).

Brasil (2019b). Proposta de emenda à constituição no. 17, de 2019. Available at: https://www25.senado.leg.br/web/atividade/materias/-/materia/135594. Accessed: 05/05/2021 (in Brazilian Portuguese).

Brasil (2021). Guia do framework de segurança. Available at: https://cutt.ly/vbRQwcn. Accessed: 05/05/2021 (in Brazilian Portuguese).

Bucci, M. P. (2008). Notas para uma metodologia jurídica de análise de políticas públicas. *Políticas públicas: possibilidades e limites*.

Can, F., Özyer, T., and Polat, F. (2014). *State of the Art Applications of Social Network Analysis*. Springer International Publishing, New York, NY.

Carvalho, L. P., Oliveira, J., Cappelli, C., and Majer, V. (2019). Desafios de transparência pela lei geral de proteção de dados pessoais. In *Anais do VII Workshop de Transparência em Sistemas*, pages 21–30, Porto Alegre, RS. SBC. (in Brazilian Portuguese).

Carvalho, L. P., Oliveira, J., and Santoro, F. (2020). Who watches you? an allegory of dataveillance and cyberstalking. In *Anais do I Workshop sobre as Implicações da Computação na Sociedade*, pages 85–96, Porto Alegre, RS, Brasil. SBC. (in Brazilian Portuguese).

Chassang, G. (2017). The impact of the EU general data protection regulation on scientific research. *ecancer*, 11(709).

Donahue, L. (2016). Who owns the content posted on social media? Available at: cutt.ly/CtmXBCR. Accessed: 05/05/2021 (in Brazilian Portuguese).

EU (2018). General data protection regulation (gdpr): Regulation (eu) 2016/679. Available at: https://gdprinfo.eu/. Accessed: 05/05/2021.

Fabiano, N. (2020). The value of personal data is the data protection and privacy preliminary condition: Synthetic human profiles on the web and ethics. In *Proceedings of the 3rd International Conference on Applications of Intelligent Systems*, APPIS 2020, New York, NY, USA. Association for Computing Machinery.

Fleck, G. (2020). Anti-fascist movements are re-emerging in brazil to counter bolsonaro. Available at: https://cutt.ly/LbRQWat. Accessed: 05/05/2021.

Gabdulhakov, R. (2020). (con)trolling the web: Social media user arrests, state-supported vigilantism and citizen counter-forces in russia. *Global Crime*, 21(3-4):283–305.

Guess, A., Nagler, J., and Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on facebook. *American Association for the Advancement of Science*, 5(1).

Harris, J. (2015). What's in a name? a method for extracting information about ethnicity from names. *Political Analysis*, 23(2):212–224.

Hijmans, H. and Raab, C. (2018). *Ethical dimensions of the GDPR*. Edward Elgar.

Isaak, J. and Hanna, M. J. (2018). User data privacy: Facebook, cambridge analytica, and privacy protection. *Computer*, 51(8):56–59.

Jernigan, C. and Mistree, B. (2009). Gaydar: Facebook friendships expose sexual orientation. *First Monday*, 14(10).

Kemeny, R. (2020). Brazil is sliding into techno-authoritarianism. Available at: https://www.technologyreview.com/2020/08/19/1007094/brazil-bolsonaro-data-privacy-cadastro-base/. Accessed: 05/05/2021.

Kotsios, A., Magnani, M., Vega, D., Rossi, L., and Shklovski, I. (2019). An analysis of the consequences of the general data protection regulation on social network research. *Trans. Soc. Comput.*, 2(3).

Larson, D. and Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5):700 – 710.

Lobo, H. (2015). O que é vício de consentimento? Available at: cutt.ly/1tmX1Qs. Accessed: 05/05/2021 (in Brazilian Portuguese).

Manzoor, A. (2017). *Ethics of Social Media Research*, pages 225–238. IGI Blobal, New York, NY.

Mello, M. T. L. and Araújo, T. F. (2015). *Eficácia e efetividade da lei de acesso à informação brasileira – uma abordagem metodológica interdisciplinar*, pages 193–196. Educs, RS, Caxias do Sul.

Moor, J. H. (2005). Why we need better ethics for emerging technologies. *Ethics Inf Technol*, 7:111–119.

Moribe, G., Júnior, O., and Monteiro, R. (2019). Comparing privacy laws: Gdpr v. lgpd. Available at: cutt.ly/8tmX0ST. Accessed: 05/05/2021.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., and Crowcroft, J. (2015). Human-data interaction: The human face of the data-driven society. Available at SSRN: https://ssrn.com/abstract=2508051. Accessed: 05/05/2021.

Mulholland, C. (2020). *A LGPD e o novo marco normativo no Brasil*. Arquipélago Editorial, Porto Alegre, RS. (in Brazilian Portuguese).

Oliveira, M. I. S. and Lóscio, B. F. (2018). What is a data ecosystem? In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, number 74 in dg.o '18, New York, NY, USA. Association for Computing Machinery.

Pentland, A. (2013). The data-driven society. *Scientific American*, 309(4):78–83.

Perry, B. L., Pescosolido, B. A., and Borgatti, S. P. (2018). *Egocentric Network Analysis: Foundations, Methods, and Models*. Structural Analysis in the Social Sciences. Cambridge University Press.

Recker, J. (2013). *Scientific research in information systems: a beginner's guide*. Springer Verlag Berlin Heidelberg, New York, NY.

Reijers, W., Wright, D., Brey, P., Weber, K., Rodrigues, R., O'Sullivan, D., and Gordijn, B. (2018). Methods for practising ethics in research and innovation: A literature review, critical analysis and recommendations. *Science and Engineering Ethics*, 24.

Reis, A. B. O. (2015). *O objeto de pesquisa em ciências sociais: para além da contemplação*, pages 154–164. Educs, RS, Caxias do Sul.

Rio de Janeiro (2018). MANUAL DE BOAS PRÁTICAS E RECOMENDAÇÕES EM MÍDIA DIGITAL. Available at: `http://www.rio.rj.gov.br/dlstatic/10112/7596427/4211621/manual_boas_praticas_2018_3.pdf`. Accessed: 05/05/2021 (in Brazilian Portuguese).

Russia (2006). Federal law of 27 july 2006 n 152-fz on personal data. Available at: `https://pd.rkn.gov.ru/authority/p146/p164/`. Accessed: 05/05/2021.

Ruviaro, L. M. and Nedel, N. K. (2017). Os limites de fiscalização do servidor público: Uma análise a partir do direito a privacidade e do princípio da primazia do interesse público. In *Anais do 4º Congresso Internacional de Direito e Contemporaneidade: mídias e direitos da sociedade em rede*. (in Brazilian Portuguese).

Sherwood, J., Clark, A., and Lynas, D. (2005). *Enterprise Security Architecture: A Business-Driven Approach*. CRC Press, 1 edition.

Stattner, E. and Vidot, N. (2011). Social network analysis in epidemiology: Current trends and perspectives. pages 1 – 11.

Sumpter, D. (2018). *Outnumbered: From Facebook and Google to Fake News and Filter bubbles The Algorithms That Control Our Lives*. Bloomsbury Sigma, New York, NY.

The Economist (2017). The world's most valuable resource is no longer oil, but data. Available at: `cutt.ly/XtmX3K5`. Accessed: 05/05/2021.

The Moscow Times (2019). 'we tracked your social media, you're gay': Russian student threatened with expulsion. Available at: `cutt.ly/jtmX4cW`. Accessed: 05/05/2021.

Tyler, T. R. (2017). Methodology in legal research. *Utrecht Law Review*, 13(3):130–141.

Vadhan, S. (2017). *The Complexity of Differential Privacy*, pages 347–450. Springer International Publishing, Cham.

Valente, R. (2020). Relatório do governo separa em grupos jornalistas e influenciadores. Available at: https://noticias.uol.com.br/colunas/rubens-valente/2020/12/01/governo-bolsonaro-jornalistas-redes-sociais.htm. Accessed: 05/05/2021.

Véliz, C. (2020). *Privacy is Power: Why and How You Should Take Back Control of Your Data*. Bantam Press, London, UK.

Yang, S., Keller, F. B., and Zheng, L. (2017). *Social Network Analysis: Methods and Examples*. SAGE Publication Inc., California, CA.

Yun, R., Rosso, A. M., Maia, F., Godinho, G., and Gonzaga, R. (2020). Comparativo - gdpr, lgpd, ccpa, pci, bacen, iso. Available at: https://www.lgpdacademicooficial.com.br/materiais/2ab39155-0c1f-49b5-aa0b-bc6cae49e956. Accessed: 05/05/2021 (in Brazilian Portuguese).

Zheleva, E. and Getoor, L. (2011). *Privacy in Social Networks: A Survey*, pages 277–306. Springer US, Boston, MA.