






Identification of the Relationships Between the Stages of the Data Lifecycle and the Principles of the Brazilian General Data Protection Act

Gislaine Parra Freund^(✉) , Priscila Basto Fagundes ,
and Douglas Dyllon Jeronimo de Macedo 

Federal University of Santa Catarina/PPGCIN, Florianópolis, Brazil
gislaineparraf@gmail.com, priscila.bfagundes@gmail.com,
douglas.macedo@ufsc.br

Abstract. The purpose of this paper is to present an analysis of the relationship between the principles of the Brazilian General Data Protection Act - LGPD and the stages of the data lifecycle. An analysis was made about the objectives of each stage of the data lifecycle and the principles of the legislation. This analysis made it possible to answer the following research question: What is the relationship between each Brazilian LGPD principle and each of the stages of the data lifecycle? The data lifecycle presents the steps in which data act in a given process and relate to each other forming a chain of dependency between them. Based on the results, it was possible to observe that the data lifecycle model can be used to support and systematize the law compliance activities, since the law principles presented relations with the model stages. It was also observed that the principles of adequacy of data processing for the purposes of use and transparency are those that guide the other principles and that, besides these, the principles of safety, prevention, responsibility and accountability recommended by law, were related with all phases of the data lifecycle.

Keywords: Data lifecycle · General Data Protection Act · Data protection

1 Introduction

Technological advancement has enabled greater access to data and thus new ethical issues to be addressed within the legislation on this subject. This scenario prompted the approval, on August 14, 2018, of Brazilian Law No. 13.7091 - General Personal Data Protection Act¹ (LGPD), based on the General Data Protection Regulation² (GDPR), a standard that regulates the processing of personal data in the countries of the European Union. Brazilian LGPD is similar to GDPR and has been sanctioned to provide Brazil with competitive conditions in conducting international business, and because it is an

¹ http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm.

² <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

extraterritorial law, an adequate adequacy of organizations is necessary, regardless of its line of business.

Brazilian LGPD provides for the processing of personal data, including in digital media, by a natural person or legal entity governed by public or private law, in order to protect their fundamental rights of liberty and privacy. According to the law, activities for the processing of personal data must comply with the following principles: purpose, appropriateness, necessity, free access, data quality, transparency, security, prevention, non-discrimination or responsibility and accountability [1].

These principles must be considered by organizations to adapt their processes, systems, and services and is necessary to define an approach to support this movement.

In order to provide support to organizations and assist them in complying with the legislation, an analysis was made of the principles proposed by the law and identified in which phases proposed by [2], for the data lifecycle, need to be observed. [2] proposes in his model that the data lifecycle be divided into four phases: collection, storage, retrieval and disposal.

If, on the one hand, the data lifecycle segments the process in phases according to its acting, on the other, the principles guide the data protection treatment activities. Thus, it is understood that the analysis presented in this study, by indicating the relationships that exist in each phase, can help organizations in the adequacy activity.

This article aims to present the results of this analysis and foster new studies involving Brazilian LGPD and the data lifecycle, including research involving the proposition of models to be used in the process of adequacy of organizations that deal in some way with storage and use of personal data to the new legislation.

2 Theoretical Foundation

2.1 Brazilian General Personal Data Protection Act (LGPD)

The Constitution of the Federative Republic of Brazil of 1988 recognizes the fundamental right to privacy and freedom of expression, which consequently guarantees the privacy of people's data and their right to instrumentalize them as their own [3]. However, globalization and the development of new technologies have fostered competition between organizations regarding the security of corporate information, as well as their customers. According to [4], companies and the government are increasingly vulnerable to espionage or malicious attacks that lead to information leakage or misuse. The Personal Data Protection Act comes to broaden the scope of protection of corporate information, including in this scenario the need to protect personal data.

Following a number of data misuse scandals in Europe and the United States involving large companies such as Cambridge Analytica and Facebook, governments in many countries have been forced to regulate access, storage, use and the dissemination of personal data of its citizens.

According to the current European GDPR on data protection, in its Article 4, personal data is characterized as a set of information concerning a living person that can lead to his/her identification and which should be protected regardless of how it is stored or processed, whether technologically or manually, provided they are organized according to predefined criteria [5]. According to the author, GDPR was based on a horizontal scope perspective, applying to companies from different areas that have contact with

personal data by any means, just as it has extraterritorial coverage, inserting itself even outside the European Union.

According to [6], Brazilian LGPD was based on GDPR which gives people the right to have control over their data, and regulates business models based on the use of personal data. This regulation is the result of an adaptation of the European Data Protection Directive³ adopted in 1995, which considered commercial use of the Internet from systems using personal data.

The GDPR has extraterritorial applicability and establishes that the circulation of personal data originating and destined in non-European Union countries must follow the same standards and security principles defined in European legislation, which has made other countries, among them Brazil, adapt to these issues and also establish its internal guidelines. Thus, Brazilian LGPD was approved to provide the country with greater security and privacy in personal data, enabling its insertion and its competitiveness in the world scenario.

The Brazilian LGPD in its Art. 6, determines a set of 10 principles that should guide the activities of processing of personal data. According to [6], by including such principles, Brazilian LGPD guarantees data subjects the right to request from public and private bodies information about how their data are used, which has a deadline to comply with the data owner's request. Processing activities for personal data should comply with the following principles [1]:

I - purpose: carrying out the treatment for legitimate, specific, explicit and informed purposes to the holder, without the possibility of further treatment incompatible with those purposes;

II - adequacy: compatibility of the treatment with the purposes informed to the proprietor, according to the context of the treatment;

III - need: limitation of the treatment to the minimum necessary for the accomplishment of its purposes, with the comprehensiveness of the relevant data, proportional and not excessive in relation to the purposes of the data processing;

IV - free access: guarantee to holders, free and easy consultation about the form and duration of treatment, as well as the completeness of their personal data;

V - data quality: guarantee to the owners, accuracy, clarity, relevance and updating of the data, according to the need and for the fulfillment of the purpose of its processing;

VI - transparency: guarantee to the holders of clear, accurate and easily accessible information about the treatment and its respective agents, observing the commercial and industrial secrets;

VII - security: use of technical and administrative measures capable of protecting personal data from unauthorized access and accidental or unlawful destruction, loss, alteration, communication or dissemination;

VIII - prevention: adoption of measures to prevent the occurrence of damage due to the processing of personal data;

IX - non-discrimination: the impossibility of carrying out the treatment for illicit or abusive discriminatory purposes;

³ <https://eur-lex.europa.eu/eli/dir/1995/46/oj>.

X - Accountability: Demonstration by the agent of the adoption of effective measures capable of proving compliance with personal data protection rules, including the effectiveness of such measures.

Brazilian law follows GDPR guidelines on responsibilities for international transfer of personal data in defining the role of personal data processing agents (controller, operator and responsible for processing personal data). Likewise, rules are set regarding governance, good practices, security and data confidentiality. The Brazilian LGPD should be applicable even in the case of companies based abroad, provided that the data processing operation is performed in the national territory [7].

It is noteworthy that the law presents other directives in their content, but for the analysis that is proposed in this article, the 10 principles presented here were listed because it is understood that they guide the other items of the law.

2.2 Data Lifecycle

The data lifecycle comprises the steps in which data acts in a given process. Such steps correspond to the phases in which the data relate forming a chain of dependency between them. For [8] there are several types of lifecycles, as well as different stages associated with them, which can be evaluated from different perspectives and at different managerial levels. According to the author, for example, at the managerial activity level, the lifecycle is managed in the business process domain. At the project or system level, however, the data lifecycle is managed in the context of product or service development.

Regarding data, new research involving the theme has enabled various ways of collecting, storing and retrieving them. [2] proposes that the data lifecycle be divided into four phases, namely: collection, storage, retrieval and disposal and presents that these phases are permeated by six factors: privacy, integration, quality, copyright, dissemination and preservation as illustrated in Fig. 1.

The purpose of the collection phase is to meet the information needs and it is in it that the activities related to the initial definition of the data to be used are developed, as well as the planning of how they will be obtained, filtered and organized, identifying the structure, format and the means of description that will be used.

During the storage phase, activities related to processing, transformation, insertion, modification, migration, transmission and any action aimed at data persistence on a digital medium are performed. The focus of this step is to enable data reuse.

The recovery phase is related to the consultation and visualization of the data. Its purpose is to enable better access and use of them. In the disposal phase occurs the identification of data that is no longer needed and that can be deleted, focusing on eliminating unnecessary data.

Regarding the six factors that permeate the phases of the data lifecycle presented by [2], they are: privacy that deals with aspects that guarantee the privacy of people or institutions related to the data to be used; integration refers to the identification and use of requirements that will provide the integration of data with other data; the data quality, which is related to aspects such as origin, collection mechanisms, physical and logical integrity, among others to be considered to ensure that the data is reliable and useful; copyright is related to respect for the copyright of the data; dissemination that is linked

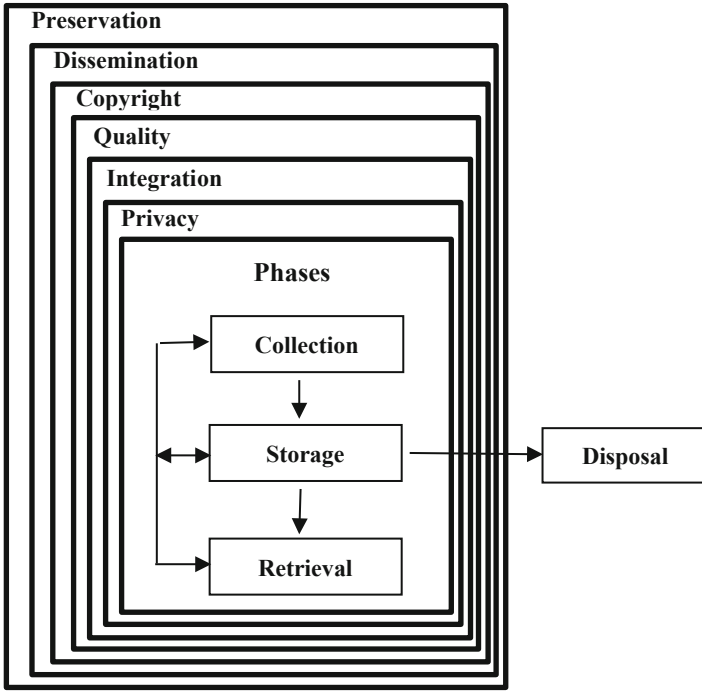


Fig. 1. Data lifecycle according to [2].

to findability and access to data; and data preservation that is related to preserving useful data so that it can be used in the future.

Table 1 presents a summary of the relationships defined and presented in [9], involving the four phases of the data lifecycle and the six factors that permeate them.

For [9], the lifecycle proposed by him aims to provide a structure that supports the efforts, studies and actions performed to obtain, maintain and use data, making it possible to approximate similar elements and distribute theories and methodologies depending on its scope, either by phases or by factors.

Table 1. Phases \times factors involved in the data lifecycle process.

	Collection	Storage	Retrieval	Disposal
Privacy	Collect while respecting privacy	Store data by adopting access control	Recover data taking users into account the content to be made available	Discard the data respecting the right of privacy and request of the “owner” of the data
Integration	Collect data that can be integrated with other databases	Store data considering how to access and adopt a SGBD ^a that enables integration with other data	Recover data with the benefits of good integration that provides greater value in its use	Discard the data by observing the consequences of deletion in relation to content derived from integrations made with other data
Quality	Collect data considering the origin and the collection mechanisms used	Store data for its physical and logical integrity	Recover data with the same quality aspects as in the collection and storage steps	Discard the data by logging information about the deletion process so as not to impair the quality of future queries and usage
Copyright	Collect respecting the copyright of the author	Store the data by linking the source of data collection	Retrieve data by explaining usage permissions and how data can be used	Discard the data while maintaining information of its own to ensure legal compliance in derivative and/or referenced works of the discarded data
Dissemination	Collect data to support data findability and access	Store data by providing interpretable and easily accessible means of accessibility	Retrieve data considering elements and strategies that allow its location and access through collection processes	Discard data by looking at the impact of eliminating key search elements and finding datasets
Preservation	Collect data so that it can be preserved	Store with the premise that data can be interpreted in the future	Recover data with the possibility of obtaining the same interpretation at different times	Discard data from systems considering keeping a copy of deleted data to preserve it if there are unforeseen demands that require it

^aSGBD - Database Management System.

3 Methodology

This study is considered a basic research as to nature, as it aims to generate new knowledge, useful for the advancement of science, with no practical application foreseen [10] and bibliographic as to procedures, which in the definition of [11], “is made from the survey of theoretical references already analyzed, and published by written and electronic means, such as books, scientific articles, web site pages”.

As for the objectives, it is exploratory, as it aims to gain a better understanding of the problem to be studied and promote greater familiarity with the themes, to make them more explicit or build hypotheses [12]. It has a qualitative approach, since it is concerned with deepening the understanding of a social group, an organization, among others without considering numerical representativeness [10].

For the analysis of the relationships between Brazilian LGPD principles and data lifecycle stages, we used the data lifecycle model proposed in [2] and Brazilian LGPD principles for data processing. personal data present in the legislation. To identify such relationships, the following question was answered: What is the relationship between each of the Brazilian LGPD principles and each of the stages of the data lifecycle proposed by [2]?

The relationships were discovered considering that Brazilian LGPD principles can have the following relationships:

- Very Relevant: for situations where the principle directly influences the data lifecycle stage - represented by the symbol ●;
- Relevant: for situations where the principle indirectly influences the data lifecycle stage - represented by the symbol ○;
- Non-Identified/Unidentified: for situations where no relationships were detected - represented by the acronym NI.

The results of the analysis will be presented at the next session.

4 Analysis of the Brazilian LGPD Principles in Relation to the Stages of the Data Lifecycle

Table 2 shows the relationships identified and then the justifications for each of them are presented. The table outlines the Brazilian LGPD principles and the data lifecycle steps in the columns.

Table 2. Relationship between Brazilian LGPD principles and data lifecycle steps.

	Collection	Storage	Retrieval	Disposal
Goal	●	○	●	NI
Adequacy	●	●	●	●
Need	●	○	●	○
Free access	NI	●	●	NI
Data Quality	●	●	●	NI
Transparency	●	●	●	●
Safety	●	●	●	●
Prevention	●	●	●	●
Non-discrimination	NI	NI	NI	NI
Accountability	●	●	●	●

● Very Relevant: Directly Influences the Data Lifecycle Step.
 ○ Relevant: Indirectly Influences the Data Lifecycle Step.
 NI: Non-Identified Relation.

The analysis identified that the **purpose** principle has a very relevant relationship with the *collection* and *recovery* stages. Considering that the *collection* has the goal of supplying the informational need, identifying the **purpose** of using the data is paramount to meet this objective. It was also identified that the **purpose** principle has a relevant relationship with the *storage* step, considering that it is necessary to store the necessary data for the identified **purpose**.

For the **adequacy** principle, a very relevant relationship was identified with all phases of the data lifecycle. Considering that this principle concerns the treatment of data appropriate to its purpose, it was observed that all phases have a direct and important relationship with this principle.

For the principle of **necessity**, a very relevant relationship was observed with the *collection* and *recovery* phases of the data lifecycle. Because it is a principle that recommends limiting treatment to the minimum necessary to achieve its purposes, it is observed that minimization in *collection* and *recovery* can be adopted. A relevant relationship with the *storage* and *disposal* phases has also been identified for this principle, considering that minimal data retention can be adopted and disposal should take into account the retention time required and consistent with the intended use of the data.

Regarding the principle of **free access**, a very relevant relationship with the *storage* and *retrieval* phases was identified, as these phases include access control in data storage, but allowing access and retrieval of the contents to be made available. For this principle, no relationships were identified with the collection and disposal phases.

Regarding the **data quality** principle, very relevant relationships were identified with the *collection*, *storage* and *retrieval* phases. The similarity of the quality factor acting in all phases of the data lifecycle with the principle of data quality favors the direct relationship identified: in the collection phase by defining elements that enable the perception of the quality of the collected data; storage phase definitions to ensure data maintains its physical and logical integrity; and the resources made available on recovery that should reflect these same aspects as described for collection and storage.

For the principle of **transparency**, a very relevant relationship was identified with all phases of the data lifecycle. Because this principle aims to provide holders with the guarantee of clear, accurate and easily accessible information on the processing of their data, this principle is directly associated with all phases, influencing the resources adopted in each of them.

By assessing the **security** principle, it was also possible to identify a very relevant relationship with the data lifecycle. It is noted that in order to protect personal data from unauthorized access and from accidental or unlawful situations concerning destruction, loss, alteration, communication or dissemination, technical and administrative measures must act from data collection to its disposal.

As with the principles of transparency and security, as regards the principle of **prevention**, a very relevant relationship with all stages of the data lifecycle has also been identified. It is considered that to prevent damage to personal data due to processing, it is necessary to adopt preventive measures at all stages since data damage can occur in any of them.

For the principle of **non-discrimination** no relationship was identified with any phase of the data lifecycle. It is understood that this principle is related to the definitions of the data use process and not to its lifecycle.

Regarding the principle of **accountability**, there was a relationship with all phases of the data lifecycle because to demonstrate the adoption of effective measures capable of proving compliance and compliance with personal data protection standards, these must be employed, practiced and validated at all stages of the cycle.

To exemplify the application approach idealized in this work, was considered the systems development process and the data collection phase. In this phase, the relevant relations with the principles of Purpose, Adequacy, and Necessity were identified. One of the alternatives to be adopted is about the principle of “collecting the minimum necessary” guided by questions, such as: What is the data really needed for system scope? What is the purpose of data collection? What is the proper way to collect such data, considering its purpose?

Thus, it is observed that the relationships presented in this study and the reflections they provoke can support organizations in the task of adaptation to the Brazilian LGPD guided by the performance of the data.

The relationships identified in this analysis are limited to indicating in the phases of the data lifecycle which principles of the law to be addressed, and their purpose is not to present what should be done or how to implement the principles of the Brazilian LGPD. The analyzes presented can be used by organizations of any nature that aim to comply with the law or assess their compliance with the law, identifying at each stage of the data lifecycle which principles need to be implemented or need to be improved.

5 Conclusions

LGPD has been sanctioned in Brazil since August 2018 and regulates the protection of personal data, ensuring the exercise of personality rights, and sets limits on the right of access to third party information and the use of such data for discriminatory, unlawful or illegal purposes. The purpose of this study was to analyze the new law, considering that its guidelines ensure the effective protection of personal data, that organizations must adapt to it within a period of 1 year after its publication and to verify if this adequacy can be guided by data lifecycle steps.

As final considerations it can be concluded that the data lifecycle model proposed by [2] and the relations existing with the principles of the Brazilian LGPD can be used as a support tool for the process of compliance by organizations.

It can also be observed that the principles of adequacy, transparency, security, prevention, and accountability are the principles that have been related to all phases of the data lifecycle, but when evaluating the purpose of each principle, the Transparency and adequacy guide the others. Considering that transparency is one of the main goals of Brazilian LGPD, this and the adequacy of data processing for use purposes, direct the other principles.

As a suggestion of future work, it is proposed that a study be conducted to identify the actions that can be taken to adapt Brazilian LGPD principles at each phase of the data lifecycle considering the relations identified with the principles of the legislation.

References

1. Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília (2018). Accessed 15 Sept 2019
2. Sant'ana, R.: Ciclo de vida dos dados e o papel da ciência da informação. In: Encontro Nacional de Pesquisa em Ciência da Informação, 14. UFSC, Florianópolis (2013)
3. Carvalho, L., et al.: Desafios de Transparência pela Lei Geral de Proteção de Dados Pessoais. In: Anais do VII Workshop de Transparência em Sistemas, pp. 21–30. SBC, Belem (2019)
4. Rapôso, C., et al.: LGPD - Lei Geral de Proteção de Dados Pessoais em Tecnologia da Informação: Revisão Sistemática. RACE Revista da Administração **4**, 58–67 (2019)
5. Costa, M.: A Era da Vigilância no Ciberespaço e os Impactos da Nova Lei Geral de Proteção de Dados Pessoais no Brasil: Reflexos no direito à privacidade. 2018. 93 f. TCC (Graduação) - Curso de Direito, Faculdade Nacional de Direito, Universidade Federal do Rio de Janeiro, Rio de Janeiro (2018)
6. Monteiro, R.: Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil? Instituto Igarapé, Artigo Estratégico 39. Rio de Janeiro, Brasil (2018)
7. Ferreira, D., Marques, R., Natale, A.: A política de informação na arena da privacidade dos dados pessoais. In: Encontro Nacional de Pesquisa em Ciência da Informação, 19. UEL, Londrina (2018)
8. Hernon, P.: Information lifecycle: its place in the management of US government information resources. *Gov. Inf. Q.* **11**(2), 143–170 (1994)
9. Sant'ana, R.: Ciclo de vida dos dados: Uma perspectiva a partir da Ciência da Informação. *Revista Informação Informação* **21**(2), 116–142 (2016)
10. Gerhardt, T., Silveira, D.: Métodos de pesquisa. UFRGS, Porto Alegre (2009)
11. Fonseca, J.: Metodologia da pesquisa científica. UEC, Fortaleza (2002)
12. Gil, A.: Como elaborar projetos de pesquisa, 3rd edn. Atlas, São Paulo (1991)