Journal of the Brazilian Computer Society, 2026, 32:1, doi: 10.5753/jbcs.2026.6448

An approach to Data Literacy through a Personalized Interactive LGPD Guide using LLM for Educators

César Murilo da Silva Junior [**Federal University of Uberlândia** | **Serviço Federal de Processamento de Dados** | ***cesar.junior@serpro.gov.br***]
Silvio E. Quincozes [**Federal University of Unipampa** | ***silvoquincozes@unipampa.edu.br***]
Juliana Saraiva [**Federal University of Paraíba** | ***julianajags@dcx.ufpb.br***]
Rafael D. Araújo [**Federal University of Uberlândia** | ***rafael.araujo@ufu.br***]

Faculty of Computing, Federal University of Uberlândia (UFU), Av. João Naves de Ávila, 2121, Bloco 1B, Santa Mônica, Uberlândia, MG, 38400-902, Brazil.

Received: 27 August 2025 • **Accepted:** 13 November 2025 • **Published:** 25 March 2026

Abstract. The Brazilian General Data Protection Law (LGPD) was created to protect the fundamental rights of freedom and privacy of Brazilian citizens. Since its implementation, it has brought new challenges to all institutions established in Brazil, whether public or private, requiring an adaptation to personal data processing practices. In the context of higher education, many professors face difficulties in understanding and properly applying the guidelines of this legislation in their daily activities. This work proposes the development of an approach to data literacy through an interactive guide, based on practical scenarios, to support educators in the process of complying with the LGPD. The proposed system uses the OpenAI API to offer personalized support in real-time. Ten representative academic scenarios were implemented, in which users can interact through multiple-choice questions followed by a chat with the guide. The results showed that, despite initial usability limitations, the system represents a promising tool to promote the comprehension of LGPD among teachers. We observed that our approach can facilitate compliance with the legislation, but requires accessibility and usability improvements to ensure greater and easier adoption.

Keywords: Data protection, Interactive Guide, Academic Institutions, Teacher, LGPD

1 Introduction

In recent years, the Brazilian General Data Protection Law (LGPD) [Brasil, 2018] has become a milestone in personal data protection in Brazil, affecting various sectors. In the education sector, the advancement of digitalization in academic processes and the growing adoption of educational technologies, ensuring compliance with the LGPD has become a priority for educational institutions [ANPD, 2023]. The legislation guarantees data subjects the right to manage their personal information, requiring institutions to adopt appropriate practices to ensure safe and transparent data handling.

Before the enactment of the LGPD, data breaches in universities had already indicated the vulnerability of the sector. An example was the accidental sending of 800 emails containing confidential student information at a university in Rio Grande do Sul [Rohr, 2021]. Even after the law was implemented, incidents continued to occur, such as a security breach on a university portal in São Paulo that allowed unauthorized access to students' personal data [Feitosa Jr., 2020]. Since August 2021, when the LGPD sanctions began to be enforced, educational institutions have faced penalties that may include significant fines and even suspension of data processing activities.

Adapting to the LGPD presents significant challenges, especially in the educational context. In addition to institutional requirements, educators also need to understand how to correctly apply the principles of the LGPD in their teaching practices. Issues such as the use of student data in assessments, interactions on virtual platforms, and sharing academic information still raise questions. Studies conducted in Brazilian educational institutions indicate limited institutional LGPD maturity and a lack of systematic training for teachers and administrators [Oliveira da Silva and Costa Sarkis, 2024; Wendling *et al*., 2023; Souza *et al*., 2022]. Although the National Data Protection Authority (ANPD) has issued guidelines for academic data processing, these materials are often static and fail to provide contextualized support for the specific realities of teaching and learning [ANPD, 2023].

International literature points to similar challenges under the European Union's General Data Protection Regulation (EU's GDPR) [European Union, 2016]. Despite longer experience with data protection frameworks, studies show that educators often lack the necessary understanding of privacy principles and compliance obligations in digital environments [Hoel and Chen, 2018; Jekabsone, 2023]. Research involving pre-service teachers also highlights insufficient preparation to handle students' personal data responsibly [Torres-Hernández and Gallego-Arrufat, 2023; Marín *et al*., 2023]. Moreover, educational data literacy frameworks emphasize that teachers must develop competencies not only in data interpretation, but also in ethical, legal, and pedagogical aspects of data use [Papamitsiou *et al*., 2021]. These findings suggest a global educational gap between the existence of regulatory frameworks and their effective implementation in daily teaching practice.

To address these challenges, this study proposes the development and evaluation of an interactive guide based on a

Large Language Model (LLM), aimed at enhancing educators' data literacy and supporting the effective implementation of the LGPD in educational settings. In this context, data literacy refers to the ability of educators to understand, interpret, and correctly apply data protection principles within the educational environment. The guide provides dynamic and context-sensitive support to help educators make informed decisions about the use and handling of personal data in various academic situations.

Unlike static materials, the utilization of an LLM allows responses to be dynamically adjusted based on user interactions, ensuring alignment with the latest guidelines from the Brazilian National Data Protection Authority (ANPD) and emerging relevant jurisprudence. This innovative approach aims not only to clarify doubts, but also to promote continuous learning tailored to the individual needs of educators.

To provide a comprehensive understanding of this initiative, the article is organized as follows: Section 2 presents the theoretical foundations related to the LGPD and data literacy in an educational context. Section 3 describes the methodology adopted for the development of the proposed solution. Section 4 details the interactive guide proposal based on LLM. Section 5 discusses the results obtained and the practical implications of the implementation. Finally, Section 6 presents the conclusions and suggestions for future work.

2 Background

2.1 Fundamental Aspects of the LGPD

The Law No. 13.709/2018 (known as LGPD) [Brasil, 2018], enacted on August 14, 2018 and inspired by the EU's GDPR [European Union, 2016], became effective on August 16, 2020, with administrative sanctions applicable from August 2021. On February 10, 2022, Constitutional Amendment No. 115 enshrined personal data protection as a fundamental right in Article 5, item LXXIX of the Federal Constitution, reinforcing the LGPD as both a regulatory framework and a constitutional duty [Brasil, 2022].

Both GDPR and LGPD regulate personal data processing in digital and physical formats, which means that even information in printed documents falls under their scope [Brasil, 2018]. In Brazil, data protection began with the Brazilian Civil Rights Framework for the Internet (in Portuguese, *Marco Civil da Internet*) (Law No. 12.965/2014) [Brasil, 2014], which established privacy principles online and paved the way for the LGPD's broader regulation. A clear example appears in its guiding principles:

> "*Art. 3 The discipline of internet use in Brazil has the following principles: (...)*
> *II - protection of privacy;*
> *III - protection of personal data, under the law*;"
> [Brasil, 2014]

Moreover, Article 7, item X of the Brazilian Civil Rights Framework for the Internet regarding data deletion was later amended by the LGPD [Brasil, 2018]. The LGPD also aligned Brazil with international standards, particularly the EU's GDPR, not only to strengthen privacy protection but also to meet economic and trade demands, ensuring compatibility for commercial relations with the European bloc [Barile da Silveira *et al*., 2024; Pinheiro, 2023].

The LGPD was a key step in aligning Brazil with international standards, particularly the EU's GDPR, driven by both privacy concerns and economic needs to sustain trade with the European bloc [Barile da Silveira *et al*., 2024; Pinheiro, 2023]. It regulates three categories of data: personal (e.g., name, date of birth, RG – Brazilian state-issued identity card – and CPF – Brazilian individual taxpayer registry), sensitive (e.g., racial origin, health, sexual orientation), and anonymized, the latter being excluded from its scope [Bioni, 2021; Doneda, 2019]. Particularly, the LGPD does not apply to the processing of personal (non-anonymized) data for the following purposes:

1. Exclusively personal and non-economic;
2. Journalistic and artistic;
3. Academic, with only Articles 7 and 11 applying;
4. Exclusively for public security, national defense, state security, or the investigation and repression of criminal offenses.

Academic purposes differ from contractual or administrative purposes set by educational institutions for students, staff, or professors [ANPD, 2022]. In research contexts, the relevant provisions are Article 7, which establishes requirements for processing personal data, and Article 11, which governs sensitive data. Specifically, Article 7(IV) authorizes its processing for studies by research bodies, provided that data are anonymized whenever possible Brasil [2018].

If the purpose is exclusively academic, provisions such as Article 8 (consent) and Article 15 (termination of processing) do not apply. Article 16 explicitly allows data retention for research by official bodies, provided the data are anonymized [ANPD, 2023]. For contractual or administrative purposes, however, the full LGPD applies, with enforcement by the ANPD, a federal agency under the Ministry of Justice and Public Security [Doneda, 2019].

In this scenario, an important actor in the LGPD implementation is the DPO (Data Protection Officer)[Brasil, 2018], responsible for developing measures to ensure the application of the law across all departments. The Information Technology (IT) department (or a similarly named unit) is often tasked with data mapping; however, some companies prefer to create a dedicated department for this purpose, while others hire a certified consultancy to perform the task.

Finally, but not less importantly, the ANPD may, at any time, request that the institution present the Personal Data Protection Impact Report. This document must describe the personal data processing procedures adopted, as well as the protection and risk prevention measures implemented [Bioni, 2021]. If any failure is identified, or an incident comes to the attention of the ANPD, it may impose the following administrative sanctions on the data controller, as specified in Article 52 of the Law:

1. Warning, with a deadline for correction;
2. Simple fine (up to 2% of revenue, capped at R$50 million per violation);
3. Daily fine;

4. Public disclosure of the violation after investigation and confirmation;
5. Blocking of personal data or partial suspension of the database;
6. Suspension or prohibition of data processing activities.

2.2 Fundamental Aspects of LLMs

Artificial Intelligence (AI) seeks to build systems capable of performing tasks that require human intelligence, such as pattern recognition, decision-making, and natural language processing. A key subfield is Machine Learning (ML), which develops algorithms that learn from data without explicit programming [Jordan and Mitchell, 2015]. Among ML approaches, deep learning stands out by using neural networks to model complex patterns in large datasets, with LLMs representing a major advance in natural language processing. Trained on massive text corpora, LLMs aim to generate coherent and contextual language, supporting tasks such as question answering, translation, summarization, and content creation. Their versatility stems from Transformer-based architectures, enabling efficient and scalable processing of natural language [Brown *et al*., 2020; Vaswani *et al*., 2017]. When considering the LGPD, LLMs offer numerous application possibilities, including:

1. **Education and training:** Providing personalized and explanatory materials for teachers regarding the requirements and implications of the LGPD.
2. **Legal assistance:** Interpreting legal clauses, explaining complex legal concepts, and suggesting corrective actions.
3. **Task automation:** Generating documents, automated responses to frequently asked questions, and practical compliance guidance.

A major advantage of LLMs is their adaptability through fine-tuning, allowing training with LGPD-specific content and pedagogical practices to provide accurate, teacher-oriented responses [Bubeck *et al*., 2023]. However, LLMs also have limitations, such as producing errors, reflecting biases, or exposing sensitive data, which demand control and validation mechanisms [Bommasani *et al*., 2021].

By combining these technologies with the interactive guide, the goal is to offer a practical and accessible tool that helps teachers meet the requirements of the LGPD while also promoting ethical and efficient management of their students' personal data.

2.3 Prompt Engineering and Personalization

Prompt engineering is the strategic design of instructions to optimize LLM outputs. Rather than issuing generic commands, it involves techniques to structure inputs for more accurate and context-aligned responses. As noted by Amatriain [2024], it goes beyond writing prompts, requiring awareness of model capabilities and constraints, and adapting instructions to each use case. Key approaches include:

- **Detailed contextualization:** Adding specific information to the prompt produces more context-aligned responses and improves accuracy and applicability, as shown by Brown *et al*. [2020].
- **Prompt priming:** Providing relevant context before interaction creates a "readiness state" that guides the model to generate more thematically aligned responses.
- **Constrained prompting:** Setting clear limits (e.g., number of sentences or LGPD compliance) ensures clarity, focus, and objectivity.
- **Prompt iteration:** Refining prompts iteratively (*e.g.,* clarity and goal alignment, then adjust to improve response quality and relevance).
- **Prompt debugging:** Calibrating difficulty and precision by carefully reviewing outputs allows for more targeted reformulation of prompts.
- **Temperature control:** Adjusting the temperature hyperparameter regulates response diversity, from deterministic/accurate (low values) to more variable/creative outputs (high values) Brown *et al*. [2020].

The application of each of these concepts in our work will be detailed in Section 5 (*Proposed Approach*).

3 Related Work

In this section, studies exploring the implementation of the LGPD in higher education institutions will be presented, as well as tools developed to support this adaptation process. The first study was carried out through a literature review using reliable sources such as Scopus, IEEE Xplore, SBC, ACM Digital Library, SSRN, Dialnet, Springer, and EBSCO-Host, which are widely recognized in the academic field. The search criteria included the string 'LGPD and (Educação or Education or School or Learning or Aprendizagem or LMS)' and a time frame from 2018 to 2024, due to the enactment of the General Data Protection Law in 2018. Given the limited number of articles, the search was adjusted to include only the string 'LGPD', limited to peer-reviewed articles published in journals or at scientific conferences. The second study, in turn, was conducted through a more general search based on the reading of websites and online resources, without adopting a methodology as rigorous as that of the first study. Although it was a more exploratory investigation based on easily accessible sources, it still aimed to cover relevant information on tools developed to support the implementation of the LGPD.

In addition to studies focused specifically on LGPD implementation in higher education institutions, there are contributions addressing other dimensions of compliance, offering a more comprehensive overview of how the topic is being tackled across different fronts. For instance, some works target *competencies for LGPD DPO roles and curricular alignment*, such as [Martins *et al*., 2025], while others propose *maturity models for governance and data compliance* in information systems institutions [Saraiva *et al*., 2025]. Meanwhile, there is research focusing on *data protection competencies in software education*, as in Saraiva *et al*. [2024], and experimental work on *LGPD artifacts and documents in agile engineering*, exemplified by Saraiva and Soares [2023]. Together, these studies illustrate that LGPD-related research spans multiple spheres—DPO training, organizational maturity, pedagogical

innovation, and engineering tools—even if not all are directly comparable, they enrich the landscape of how educational institutions engage with data protection.

3.1 Studies on the Implementation of the LGPD in Higher Education Institutions.

The selected works for this study, published between 2019 and 2024, with a higher concentration in 2023, address the implementation of the LGPD in higher education institutions. Ten of these studies broadly discuss the adaptation of privacy and security policies, the training of the involved parties, and the processing of personal data. In addition, five focus specifically on student data protection, while one analyzes the processing of teachers' data. Table 1 summarizes the thematic categories addressed, highlighting how the different studies and this dissertation relate to the areas of institutional compliance, data processing, and training in accordance with the LGPD.

Table 1. Thematic categories addressed indicating the focus of the selected works and this dissertation.

Work	C1	C2	C3	C4
[Souza *et al*., 2020]	X			
[Souza *et al*., 2022]	X			
[Gomes *et al*., 2023]		X		
[de Almeida *et al*., 2023]	X			
[Reis *et al*., 2024]	X			
[Rojas, 2020]	X			
[Stelzer *et al*., 2019]	X			
[Teodoro *et al*., 2023]	X			
[dos Santos and da Silva Pereira, 2023]	X			
[Marques, 2021]	X			
[Jesus, 2022]	X			
[Silva Junior, 2020]		X		
[Zappelini *et al*., 2023]		X		
[Novais *et al*., 2024]		X		
[de Almeida *et al*., 2020]		X		
[da Silva and Takeshita, 2023]			X	
This dissertation	X			X

Legend:
C1 - Institutional Compliance Strategies with the LGPD;
C2 - Student Data Processing;
C3 - Teacher Data Processing;
C4 - Teacher Training and Awareness;

The analyzed studies reveal diverse approaches and challenges in the implementation of the LGPD in higher education institutions. Research at IFSP identified security flaws and a lack of training, pointing to the need for continuous awareness and adaptation [Souza *et al*., 2020][Souza *et al*., 2022]. FGV developed an adequacy model for the processing of student data, with specific guidelines for different categories [Gomes *et al*., 2023]. At UFMT, a methodology adapted from the Digital Government Secretariat (SGD) guide sought efficiency in the planning and execution of compliance actions, focusing on the Personal Data Inventory and the Impact Report [de Almeida *et al*., 2023]. A systematic analysis highlighted the challenges of the LGPD concerning privacy in educational environments, such as the collection of sensitive data and the integration of technologies [Reis *et al*., 2024].

Other studies explored the implementation of the LGPD at IFSC and UFSC, identifying non-conformities and the need for adequacy programs and the appointment of a DPO [Rojas, 2020; Stelzer *et al*., 2019]. UDESC used an action research methodology to develop a business process model for adaptation, considering IT and data governance [Teodoro *et al*., 2023]. At UFS, despite training initiatives, the participation of staff was low, indicating the need for greater dissemination and the offer of more courses [dos Santos and da Silva Pereira, 2023]. A research at UFRGS analyzed the application of the LGPD from the perspective of the information regime [Marques, 2021]. The University of Rio Verde analyzed a proposal for a compliance project based on ISO 27701 [Jesus, 2022].

The relationship between cyberbullying and the LGPD was discussed, focusing on the protection of children and adolescents' data, and the promotion of Digital Citizenship [Silva Junior, 2020]. Remote teaching during the pandemic also raised questions about the data protection of minors, including challenges in obtaining consent and handling sensitive data [Zappelini *et al*., 2023]. Online security in educational institutions was addressed, emphasizing the importance of security tools, policies, and digital awareness [Novais *et al*., 2024]. Ethics, security, and privacy in distance education (EAD) during the pandemic were also analyzed, highlighting gaps in data protection and the need for robust policies [de Almeida *et al*., 2020]. Finally, one study addressed the preservation of teachers' privacy in different contractual phases, pointing to the lack of consent for the processing of their data and the devaluation of the profession [da Silva and Takeshita, 2023].

3.2 Support Tools for LGPD Compliance

This section discusses technological tools available to support LGPD compliance in organizations, highlighting consent management platforms and Data Protection Impact Assessment (DPIA) tools. While crucial, these are often not directly accessible to teachers in educational settings. The development of teacher-focused systems offering practical, real-time support is proposed as a solution.

Consent management software, exemplified by OneTrust and TrustArc [OneTrust, 2023; TrustArc, 2023], helps institutions collect, manage, and document student consent transparently. DPIAs, facilitated by tools like the GDPR Compliance Toolkit [EDPB, 2023], are essential for identifying and mitigating data processing risks, especially in high-risk academic contexts.

Continuous training for teachers and staff on privacy and data protection is vital. Online programs from platforms like Coursera and edX [Coursera, 2023; edX, 2023], along with the inclusion of LGPD courses in university curricula, can foster a culture of compliance. Furthermore, data governance systems, such as the Data Governance Framework by DAMA International [DAMA International, 2023], can improve data security and transparency.

Combined, these tools aid in LGPD compliance and promote a safer, more privacy-respecting educational environment. Table 2 compares the functionalities of existing tools with the proposed teacher-focused guide, emphasizing the guide's unique aim of supporting teachers directly.

This section has covered key aspects of the LGPD, its application in higher education, and the supporting tools. The next section presents a proposed system developed to support

Table 2. Comparison of functionalities provided by existing tools and the proposed guide

Tool	F1	F2	F3	F4	F5	F6	F7
[OneTrust, 2023]	X	X	X	X	X	X	
[TrustArc, 2023]	X	X	X	X	X	X	
[EDPB, 2023]	X	X	X	X		X	
[DAMA International, 2023]	X			X	X		
LGPD Guide for Educators					X		X

Legend:
F1 - Consent and Preference Management;
F2 - Data Subject Access Requests (DSAR);
F3 - Data Mapping and Discovery;
F4 - Privacy Impact Assessments (PIA);
F5 - Integration with AI technologies;
F6 - Technological Risk and Compliance Management;
F7 - Exploration of real-life scenarios experienced by teachers.

teachers in complying with the LGPD.

4 Method

This section outlines the research and development process, as visually detailed in Figure 1. It begins with a foundational phase (literature review), requirements gathering, and scenario definition, supported by ethical approval from the institutional committee board. The development phase was divided into front- and back-end work, along with integration with an LLM and the selection of appropriate technologies. Once the system was built, it underwent software testing and educator tests. Finally, the process culminates in the evaluation and analysis of results, ensuring that both technical performance and user acceptance are systematically addressed.

4.1 Participants and Recruitment

The study participants were professors at the Federal University of Uberlândia. According to the 2024 institutional yearbook [PROPLAD/UFU, 2025], the faculty comprises 1,927 tenured professors. Recruitment was carried out in two stages. The first stage was exploratory and involved a broad invitation sent to a general list of faculty emails, aiming to identify critical usability issues and system problems. The second stage targeted direct invitations to 285 professors, sent individually to their institutional emails after system improvements, with the goal of improving engagement and collecting more detailed feedback. Each invitation included a link to access the system, which was available online.

The preliminary stage, conducted in October 2024, included 22 sessions, primarily with professors using the system for the first time. Ten sessions were empty, with no recorded interactions or instrument responses, while only seven of the remaining 12 non-empty sessions included completed responses. The second stage resulted in 47 recorded sessions, 41 of which involved first-time users of the LGPD Guide. This targeted recruitment allowed for closer communication with participants, improving both participation and the quality of feedback collected.

The study received ethical approval from the Institutional Review Board (CEP, from Portuguese, *Comitê de Ética em Pesquisas com Seres Humanos*) at UFU via Plataforma Brasil[1], ensuring that all procedures complied with ethical standards for research involving human participants.

4.2 Procedures

Following the initial planning, the research undertook a comprehensive literature review, which is detailed in the previous section. This review aimed to understand the current state of LGPD compliance in educational institutions by analyzing scientific articles, technical reports, and legal documents to identify relevant tools and practices for both institutions and educators. Ultimately, systematizing this information informed the development of the proposed "LGPD Guide for Educators", focusing on clear explanations, best practice suggestions, and interactive self-assessment tools. This phase also included the requirements gathering, identifying both functional requirements and the subsequent choice of technologies for development and LLM.

The next important step was the selection of the LLM, which involved a comparative analysis of models like GPT-3.5-Turbo-0125, GPT-4-Turbo, Gemini 1, Claude 2, and Llama 2, based on technical performance, ease of integration, cost, and institutional support. This analysis led to the choice of GPT-3.5-Turbo-0125 via the OpenAI Application Programming Interface (API)[2] due to its balance of performance, cost-effectiveness, ease of integration, and support for academic projects.

Ten practical scenarios were created based on the authors' experiences as university teachers, each with an average of 7 years. They illustrate common situations faced by higher education educators and serve as examples for reflecting the application of the LGPD in academic daily life. They range from the storage and sharing of basic personal data from students, such as contact lists and grades, to more complex situations, such as research involving sensitive data, the use of messaging groups, attendance management, and the publication of results. They also highlight dilemmas related to information sharing among professors, the role of the DPO, the actions of the ANPD, and students' rights to request the deletion of their data. In summary, the scenarios emphasize the importance of adopting appropriate technical and administrative measures—such as access restrictions, the use of institutional systems, encryption, anonymization, and consultation with the DPO—ensuring that data processing respects students' privacy and complies with current legislation.

The development of the interactive guide followed an incremental and iterative approach, which included front-end and back-end development, focusing on usability, integration with the LLM for personalized responses, and rigorous software testing with potential users to refine the system and ensure the quality of LLM-generated responses.

4.3 Instruments

As already mentioned, the evaluation employed two rounds of experiments, both guided by the Technology Acceptance

[1]Certificate of Submission for Ethical Review (CAAE, *Certificado de Apresentação para Apreciação Ética*): 78657524.4.0000.5152

[2]https://openai.com/pt-BR/api/

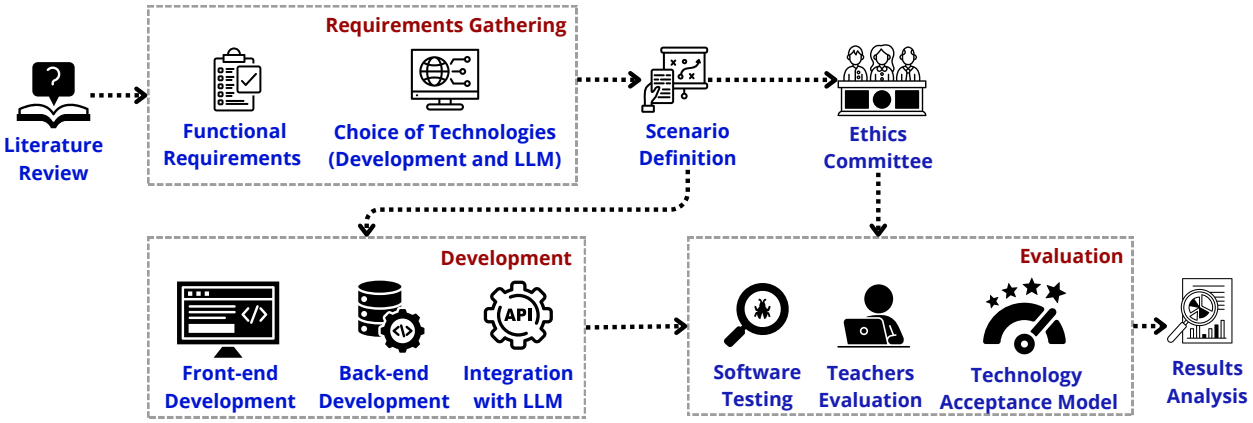

Figure 1. Flow of the methodological stages of this research.

Model (TAM) [Davis, 1985, 1989], to assess educators' perceptions of the LGPD Interactive Guide. This process involved designing a test execution plan and making the system available to educators.

The TAM is a theoretical framework widely used to predict and explain user behavior regarding new technologies, in terms of:

- **Perceived Ease of Use (PEOU):** the degree to which users believe the system will be effortless to use;
- **Perceived Usefulness (PU):** the extent to which the system is considered beneficial for task performance;
- **Perceived Intention to Use (PIU):** the intention of users to continue using the system.

There are numerous adaptations and validations of this model for specific contexts. The instrument used in this study is fundamentally grounded in the original and validated TAM [Davis, 1985], as opposed to adaptations like TAM2 [Venkatesh and Davis, 2000], because the study focuses on the individual acceptance of technology by educators, without the need to account for social or external influence factors. Since the analysis centers on the direct interaction between educators and the guide, the original TAM provided a simpler and more appropriate framework. It is important to note that, due to the small sample size, the instrument was not validated in the context of this study. However, other studies have already validated instruments based on the TAM for the Brazilian settings [Pinto *et al*., 2019; Costa Filho *et al*., 2007].

Table 3 presents the statements used in the questionnaire, grouped by dimensions. It was originally administered to participants in Portuguese. However, for the purpose of this article, the instrument's statements have been literally translated into English to ensure methodological transparency and facilitate the readers' understanding of the specific wording used in the study.

A 5-point Likert scale, ranging from "Strongly Disagree" to "Strongly Agree", was used to measure TAM constructs Likert [1932]. This method enabled teachers to express nuanced levels of agreement and provided detailed insights into perceptions of usability and usefulness.

In addition to the TAM questionnaire, the system incorporated a continuous response evaluation mechanism for the outputs generated by the LLM. After every three interactions in each scenario, a modal prompted users to rate responses using a star-based system on three criteria:

- **Clarity**: whether the response was clear and easy to understand;
- **Relevance**: whether the response addressed the user's question;
- **Usefulness**: whether the response helped solve the user's problem.

This mechanism captured immediate feedback on the quality of AI-generated responses. To ensure broader participation, even from those who did not complete the entire guide, the modal also included a TAM-related question about the clarity and comprehensibility of the interactive guide.

4.4 Data Analysis

The analysis of the evaluation data was carried out along two complementary dimensions. First, the results of the TAM questionnaire were examined. Descriptive statistics were calculated for the variables (PEOU, PU, and PIU), providing an overview of educators' perceptions of the proposed solution. In addition, the distribution of responses was analyzed for both the preliminary experiments and the subsequent evaluation phase. Boxplots were used to illustrate the dispersion of responses on the Likert scale, providing a visual representation of the medians, quartiles, and variability across participants. The sessions' average duration and the number of correct/incorrect answers in the interacted scenarios were also calculated.

Second, the evaluation of LLM-generated responses was analyzed. For each scenario, the number of evaluations was recorded, and average scores were computed for the three established criteria: Clarity, Relevance, and Usefulness. These measures offered a quantitative assessment of educators' immediate perceptions of the quality of the AI-generated content and complemented the insights obtained from the TAM questionnaire.

4.5 AI Assistance Statement

The authors acknowledge the use of ChatGPT as a language assistance tool during the manuscript preparation and revision

Table 3. TAM Questionnaire Statements.

Dimension	Statement
Perceived Ease of Use (PEOU)	My interaction with the LGPD Guide is clear and understandable.
	Interacting with the LGPD Guide requires little mental effort.
	I find the Guide easy to use.
	I find it easy to get the system to do what I want.
	It is easy to remember how to use the Guide.
Perceived Usefulness (PU)	The Guide helped me understand fundamental LGPD concepts and guidelines.
	The Guide helped me comply with the LGPD guidelines.
	Using the Guide was more practical than searching for information and clarifying doubts through other means.
Perceived Intention to Use (PIU)	Having access to the Guide, I intend to continue using it.

process. The tool was employed exclusively to enhance the clarity, coherence, and grammatical accuracy of the English text. All scientific content, analyses, interpretations, and conclusions presented in this manuscript were independently developed and verified by the authors.

5 Proposed Approach

This section presents the proposed approach, which is based on an interactive guide to promote data literacy for educators and, consequently, helps educational institutions comply with the LGPD. This system was developed with the aim of providing easy and practical access to demands related to data protection and privacy in the academic context.

5.1 System Overview

The system is designed to provide an interactive interface in which educators can ask questions and receive guidance on how to implement the LGPD in their daily activities. The structure of the system is divided into practical scenarios that reflect common situations faced by educators when dealing with students' personal data and other sensitive information.

Each scenario begins with a multiple-choice question. The system responds by indicating whether the chosen alternative is the most correct, and from there an open interaction follows, in which the user can exchange ideas with the guide. Due to the integration costs associated with LLM, the interaction for the experiments in this research was limited to a maximum of three exchanges per teacher.

5.2 Technological Infrastructure

For the implementation of this system, modern and effective technologies were chosen to ensure a smooth and responsive user experience. The main technologies used are as follows:

LLM and ChatGPT: The GPT-3.5-Turbo-0125 model was used to enable natural and contextualized interactions with users. This model is responsible for understanding the questions asked by teachers and generating accurate and informative responses about the LGPD [OpenAI, 2024].

Angular: This web development framework was chosen for its ability to create dynamic and responsive user interfaces. Angular facilitates the creation of an application that is not only visually appealing but also highly functional, allowing intuitive navigation [Angular, 2024].

Firebase Hosting and Firebase Realtime Database: Firebase was chosen as the platform for hosting and managing the system's data. Firebase Hosting provides a secure and scalable environment for the application, while Firebase Realtime Database allows for the storage and realtime synchronization of data, ensuring that users have access to updated and relevant information during their interactions [Firebase, 2024].

Git and GitHub: To ensure the security and versioning of the project, Git was used as the version control system, and GitHub was used as the platform for hosting the repositories. Using Git allows incremental changes to be recorded in the code, making it easier to track development and revert to previous versions if necessary. GitHub provides a secure and centralized copy of the project, accessible from any device with an internet connection, and allows efficient collaboration if needed [Git Development Community, 2024; GitHub, 2024].

5.3 Interface and Interaction Flow

The system interface was carefully designed to provide an intuitive and accessible experience for teachers who wish to understand and apply the LGPD guidelines in their academic environment. By organizing the interaction steps, from accepting the Free and Informed Consent Form (TCLE, from Portuguese, *Termo de Consentimento Livre e Esclarecido*) to the final evaluation of the user experience, the system aims to facilitate learning and engagement. In the following, we describe each step of the interaction flow, illustrating the main screens and features.

On the system's home screen, the teacher is presented with the TCLE, which outlines the research objectives, emphasizes the voluntary nature of participation, and ensures the privacy of the collected data, which will be analyzed in an anonymized manner. The form also provides direct contact information for researchers in case of any clarifications. To proceed, the teacher must accept the form.

Figure 2 shows the beginning of the interaction with the guide. Here, the first scenario is presented with a multiple-choice question, in which the teacher selects the option they consider most appropriate in relation to the scenario and the LGPD. This initial choice serves two main purposes: first, to offer a more guided and structured introduction, helping the

teacher become familiar with the dynamics of the guide before proceeding to more open interactions; second, to provide an initial assessment of their knowledge about the LGPD. In the end, a summary will be presented with the number of correct answers among the 10 proposed scenarios.

Both the question and its alternatives were created with the help of ChatGPT through an iterative prompt engineering process, during which successive adjustments were made to calibrate the grammar, difficulty level, and compliance with the law. The model was instructed to generate five options: one fully compliant with the LGPD and four with subtle discrepancies. To refine the responses, techniques such as detailed contextualization and prompt debugging were applied, ensuring that the incorrect options were plausible enough to challenge the teacher's critical thinking. Furthermore, the difficulty was progressively adjusted using strategic prompt rewording to avoid overly obvious answers, thereby making the experience more reflective and instructive.

As illustrated in Figure 3, when selecting an option, the teacher receives an initial response from the LLM related to that selection. From this point onward, they can continue the interaction by sending additional questions or comments to the guide, allowing for a deeper exploration and a more detailed understanding of the presented scenario.

As mentioned, the system uses an LLM to evaluate teachers' responses regarding the LGPD. Each teacher's message prompts the model to assess compliance with the LGPD, identify any violations by citing specific articles, and engage the teacher with direct questions and clear solutions, assuming no advanced knowledge. The interaction is designed to be concise, with the model's responses limited to three sentences.

Throughout the development process, the system's interaction instructions were refined using prompt engineering techniques to enhance the clarity, objectivity, and coherence of the model's responses. Strategies such as prompt priming were employed to provide context from the first interaction, while constrained prompting was used to limit responses to a specific number of sentences, ensuring brevity and focus. In addition, an iterative refinement approach was adopted, incorporating feedback from teachers to continuously improve the prompts and the overall quality of interaction. This methodological enhancement process is detailed in the Experiments Chapter, particularly in the section addressing identified challenges.

After three interactions within a scenario of the LGPD Guide, an evaluation modal is presented to the user, as shown in Figure 4. This modal appears with a 10-second delay after the guide's last response, ensuring enough time for the user to understand the response before evaluating it.

In the modal, the user can evaluate three aspects of the responses provided for the current scenario: Clarity, Relevance, and Usefulness. In addition, a question from the TAM is already presented, allowing for the collection of general perceptions regarding the experience with the guide. As explained in the methodology, in the section discussing the system and scenario evaluation, this approach aims to increase the response rate and ensure feedback even from users who do not advance to the final scenario.

Finally, the TAM questions screen (Figure 5) displays the first question of the evaluation, capturing the teacher's perceptions about the usability and usefulness of the system. Note that there is a last option "Prefer not to answer" that is not part of the 5-point Likert scale. This option was included as part of an ethical requirement and was not included in the analysis of the results.

These steps were planned to ensure a continuous and in-depth understanding of data protection practices, allowing teachers to progress from initial acceptance of the TCLE to the application of the acquired knowledge. The strategy for developing the interactive guide is presented, which outlines how this tool was structured to maximize clarity and user engagement during the guidance process.

5.4 Integration of the Guide with the LLM API: Financial, Operational, and Technical Aspects

The choice of an LLM API was preceded by research on the available options, considering both the quality and the costs per request. It was estimated that with the interaction of at least 100 teachers, the costs could be substantial, especially if each interaction generated multiple requests.

To mitigate financial impacts and ensure the continuity of the project, OpenAI requested credits through their Researcher Access Program[3], aimed at supporting studies on the responsible deployment of AI and risk mitigation. The request highlighted the development of a website in Angular, integrated with OpenAI's library, and featuring controlled interaction.

In addition, token usage control was a critical aspect of minimizing costs. A 'token' in the context of LLM refers to a unit of text processing, which could be a word, part of a word, or even a symbol, depending on the content of the interaction. Teachers' interactions were monitored to ensure that their questions stayed within the scope of the LGPD, with notifications sent in case of deviation. This approach allowed for the balance of the quality of the support provided with the financial feasibility of the project, ensuring the sustainability of the interactions.

5.5 Results

This section presents the results obtained from the analyzes in three parts: Preliminary Evaluation, Case Study of the Interactive Tool, and Evaluation of Responses for Each Scenario. The results are organized to highlight the main findings from each stage and to show how participant feedback contributed to the refinement of the Interactive Guide.

5.5.1 Preliminary Evaluation

A preliminary evaluation was conducted with an initial group of teachers, totaling 22 sessions throughout the month of October. Of these, 17 sessions were conducted by teachers who reported that it was their first time using the system. Of all of them, 10 sessions were observed to be empty as they did not register any interaction or response to the TAM (they did not evaluate the system). Of the remaining 12 non-empty

[3] https://openai.com/form/researcher-access-program/

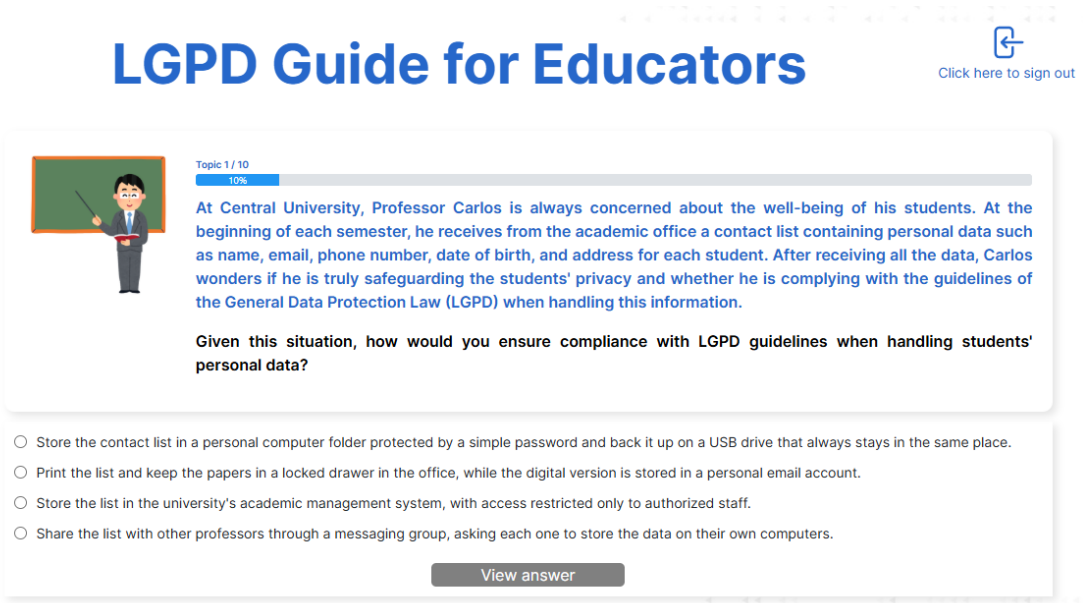

Figure 2. Example scenario and multiple choice question.

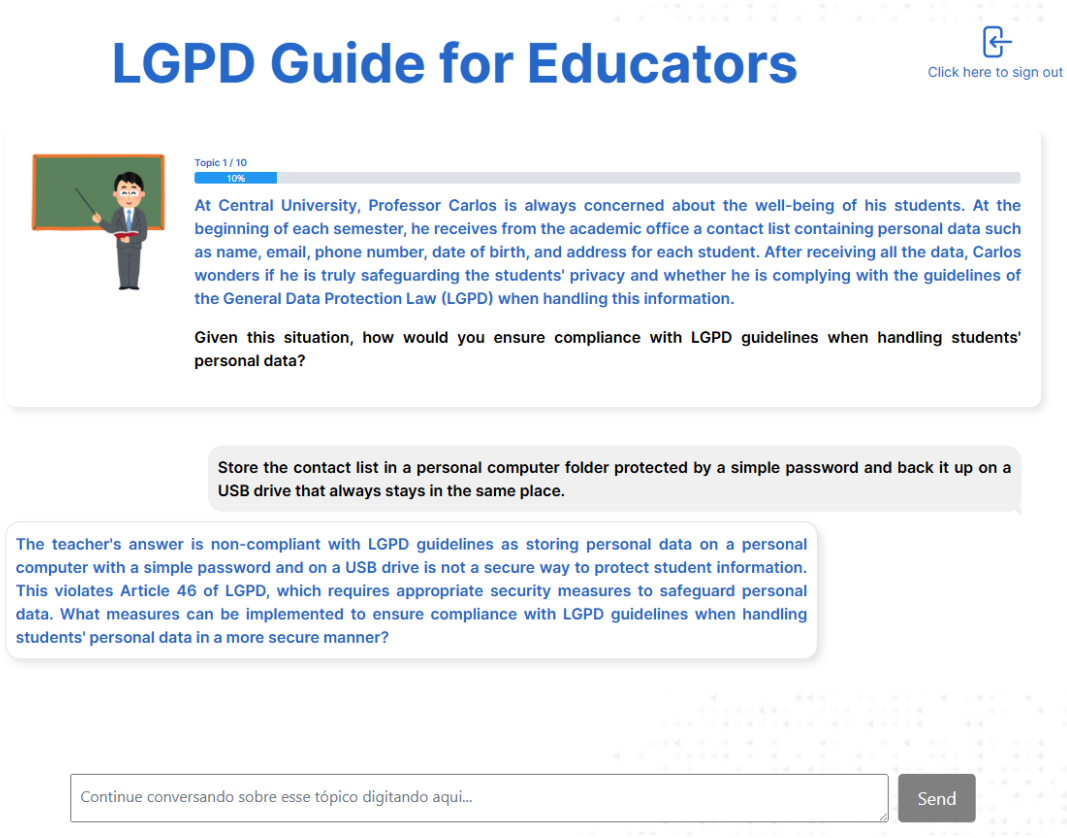

Figure 3. Guide's response to the selected alternative.

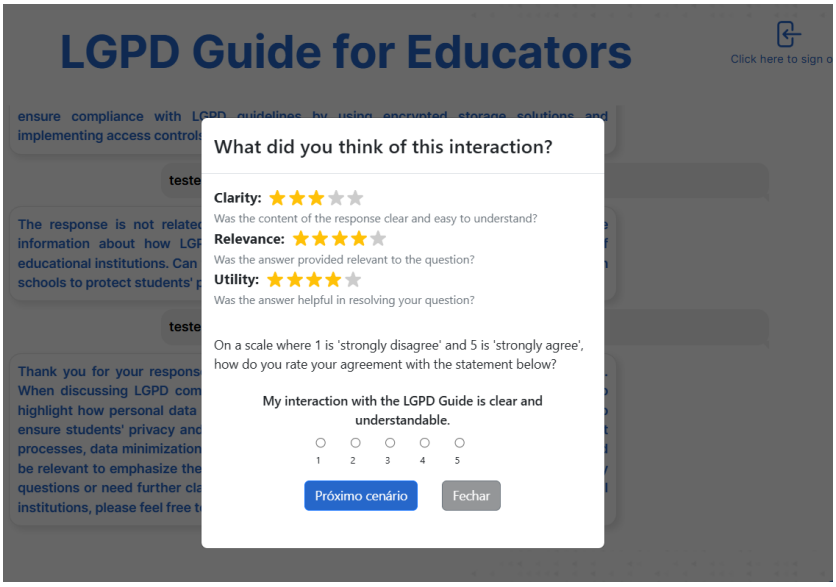

Figure 4. Evaluation of the Interaction and the Guide in general.

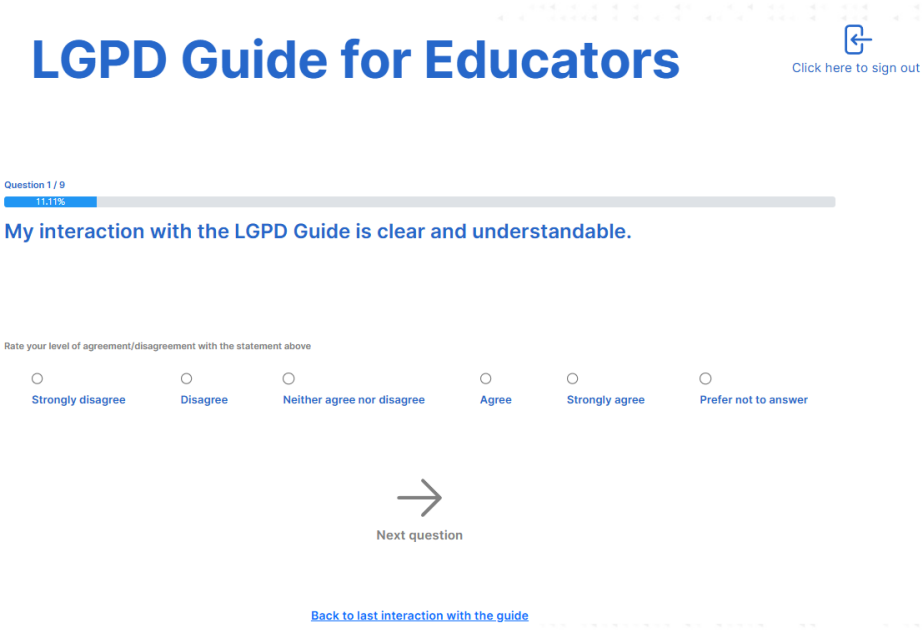

Figure 5. Screenshot of one of TAM's question.

sessions, only seven responded to the TAM. The average number of interactions per session was 1.73, with a standard deviation of 2.93, indicating a high variability in the number of interactions between sessions. This standard deviation, which is higher than the mean, reflects that interactions are not evenly distributed, and some sessions have very different numbers of interactions. Additionally, the highest number of interactions recorded in a single session was 10, which is significantly higher than the average, suggesting the presence of extreme values that contribute to this variability.

Next, Figure 6 presents response data from the TAM questionnaire, with responses distributed across the Likert scale categories of 1 to 5 (where 1 represents "Strongly Disagree" and 5 represents "Strongly Agree").

Based on these responses, Table 4 presents the descriptive statistics of the responses to the TAM questionnaire, detailing the mean, median, standard deviation, minimum, and maximum values for each of the dimensions evaluated.

Table 4. Descriptive statistics of TAM questionnaire dimensions - preliminary evaluation.

Statistic	PEOU	PIU	PU
Valid	35	7	21
Median	3.000	3.000	3.000
Mean	3.114	2.857	2.810
Standard Deviation	1.157	1.345	1.167
Minimum	1.000	1.000	1.000
Maximum	5.000	4.000	5.000

PEOU - Perceived Ease of Use;
PIU - Perceived Intention to Use;
PU - Perceived Usefulness.

Table 4 shows that PEOU has a mean slightly above 3.0, indicating a moderately positive evaluation. PIU has a slightly lower average value, suggesting that participants may have doubts about the system's continued adoption. PU is also close to 3.0, reflecting a balanced perception of the system's benefits.

Based on these statistics, the data was analyzed using a boxplot, which displays the distribution of responses for each dimension of the questionnaire: PEOU, PIU, and PU. The graph is presented in Figure 7.

The boxplot shows the dispersion of the responses on a Likert scale, allowing us to visualize the median, quartiles, and data variability. The three dimensions analyzed (PEOU, PIU, and PU) have medians equal to 3.0, indicating a generally neutral to slightly positive perception. However, the data dispersion shows that there are both lower and higher evaluations, highlighting differences in user experiences with the system. PEOU shows a slightly lower median, suggesting that some aspects of this dimension might have been more challenging for users, while PIU and PU have more balanced distributions.

5.5.2 Case Study of the Interactive Tool

Unlike the preliminary evaluation, where the invitation was made more broadly via individual institutional email. At this stage, invitations were sent directly to 285 educators. This approach aimed to establish closer and more direct communication with potential participants, resulting in a total of 47 recorded sessions, of which 41 were conducted by individuals who reported it was their first interaction with the LGPD Guide. In total, 10 sessions had no interaction. The average number of interactions per session was 15.27, with a standard deviation of 10.39. Among these sessions, 28 included responses to the TAM questionnaire, and 20 of them were complete, which means that all nine items of the instrument were answered.

Figure 8 shows the responses obtained for each evaluated question. The results indicate a generally positive perception of the LGPD Interactive Guide. Most participants agreed or strongly agreed that the guide was clear, easy to use, and required little mental effort to interact with. Educators also highlighted its usefulness in understanding key LGPD concepts and adapting to its guidelines. Furthermore, participants considered the guide more practical than seeking information through other sources and expressed an intention to continue using it in the future.

Table 5 presents the descriptive statistics of the three dimensions of the TAM questionnaire analyzed in the case study. The results show that all three dimensions have relatively high central tendency values, with medians equal to 4.000 and means ranging from 3.790 to 4.117, indicating generally positive perceptions among participants. The variability is moderate, with standard deviations between 0.825 and 1.131, suggesting some dispersion in responses, particularly for PEOU. The minimum and maximum scores demonstrate the full use of the Likert scale (1 to 5), revealing that participants expressed both low and high evaluations across items. Overall, these statistics suggest that respondents found the system easy to use, useful, and were inclined to use it in the future, with slightly higher agreement regarding perceived usefulness and intention to use.

Table 5. Descriptive statistics of the TAM questionnaire dimensions - case study.

Statistic	PEOU	PIU	PU
Valid	100	20	60
Median	4.000	4.000	4.000
Mean	3.790	4.050	4.117
Standard Deviation	1.131	0.826	0.825
Minimum	1.000	3.000	2.000
Maximum	5.000	5.000	5.000

Legend:
PEOU - Perceived Ease of Use;
PIU - Perceived Intention to Use;
PU - Perceived Usefulness.

In addition, the boxplot presented in Figure 9 provides a clear representation of the dispersion and concentration of the responses given by the participants for each TAM dimension and reflects the distribution of the data in a more informative way than simple means and standard deviations.

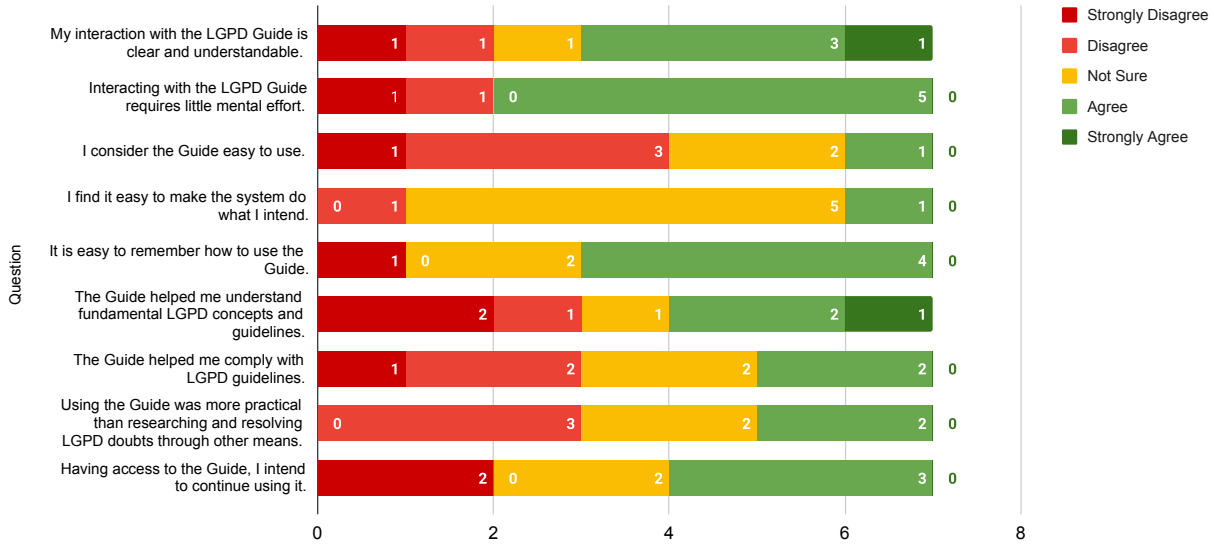

Figure 6. Distribution of participants' responses to the TAM questionnaire regarding the usability and effectiveness of the LGPD Guide, preliminary experiments.

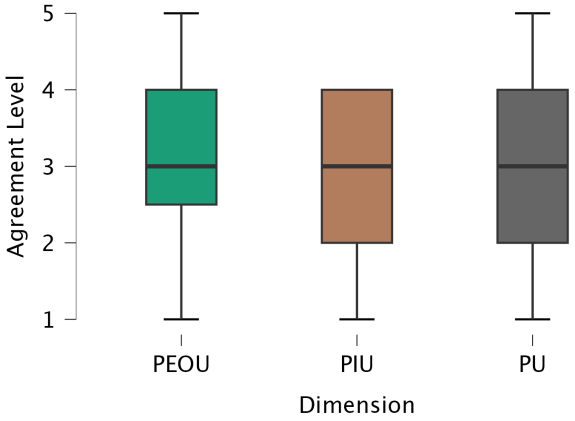

Figure 7. Distribution of TAM responses for the preliminary experiment.

Legend:
PEOU - Perceived Ease of Use;
PIU - Perceived Intention to Use;
PU - Perceived Usefulness.

5.5.3 Evaluation of Responses for Each Scenario

In addition to the general evaluation of the guide's use through the TAM, the interactions for each scenario were also evaluated. As explained, after interacting three times with each scenario, a modal is displayed after 10 seconds for teachers to evaluate the clarity, relevance, and usefulness of the responses generated by the LLM. Figure 4 shows a screenshot of the evaluation interface, where the teacher selects the number of stars corresponding to their perception of response quality (1 star indicating "very poor", 3 for "neutral", and 5 for "very good").

The evaluation of the interactions provided relevant insights into the educators' perception of the quality of support offered by the interactive guide. A total of 23 educators responded to this evaluation, but as shown in Table 6, not all evaluated all scenarios, resulting in a variation in the number of responses per criterion and scenario. It can be observed that the initial scenarios received more evaluations than the latter ones, suggesting greater attention from teachers at the beginning of their use of the guide.

Figure 10 shows the distribution of the average ratings by scenario and criterion, represented by grouped bars. Each scenario contains three bars corresponding to the average ratings for clarity, relevance, and usefulness, allowing for a comparative analysis between the different contexts presented to the teachers.

5.5.4 User Engagement and Response Accuracy

The total time spent by each participant during their interaction sessions with the platform was computed for further analysis. The system automatically logged the timestamp of each interaction, allowing for the calculation of individual and aggregate metrics. The average session time was 1 hour and 16.81 minutes, with a median of 15 minutes and 48 seconds. The shortest session lasted 37.15 seconds, while the longest reached 22 hours and 48 minutes. These values reflect the diversity of engagement levels among educators, ranging from brief exploratory interactions to prolonged and in-depth use of the guide. Very long sessions, such as the one with the longest duration, were not expected, and it may have happened that the participant left the browser open between one interaction and another to do other activities.

Among the 47 sessions in the case study, the system recorded 156 correct answers and 76 incorrect answers in all ten scenarios. This distribution suggests that, overall, participants demonstrated a reasonable level of understanding and engagement with learning activities about LGPD. However, the observed error rate (around 32%) indicates that the scenarios presented an appropriate level of challenge, providing opportunities to identify misconceptions and learning gaps.

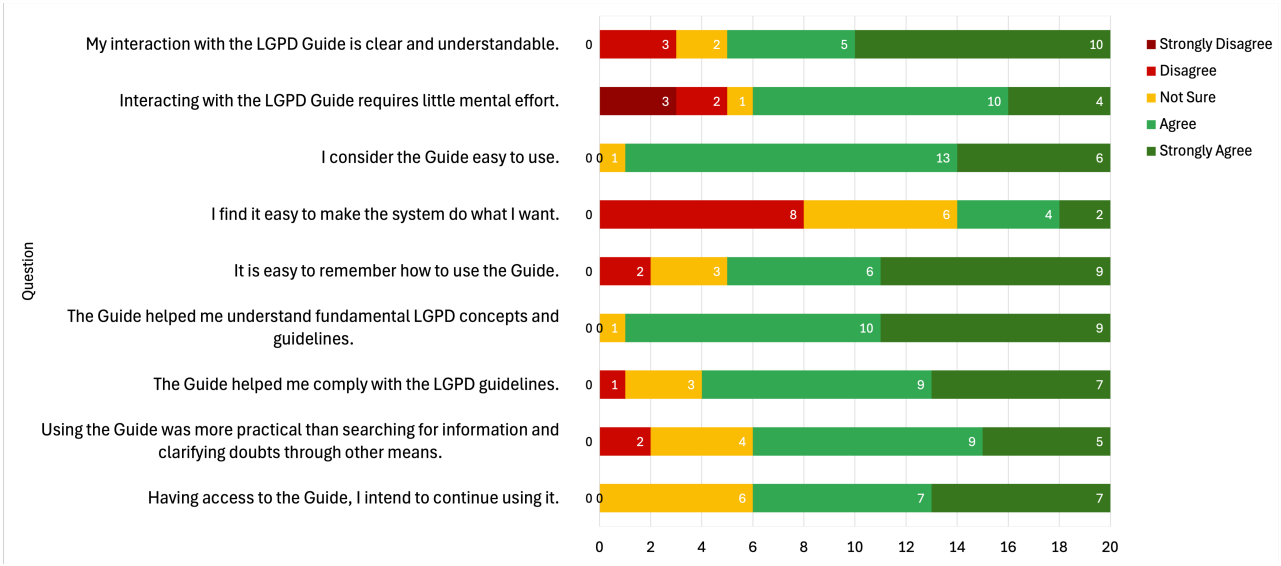

Figure 8. Distribution of participant responses to the TAM questionnaire about the usability and effectiveness of the LGPD Guide.

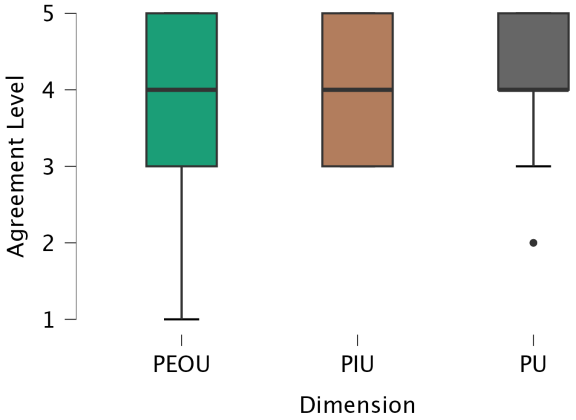

Figure 9. Distribution of TAM responses for the case study.

Legend:
PEOU - Perceived Ease of Use;
PIU - Perceived Intention to Use;
PU - Perceived Usefulness.

The analysis of response accuracy across the ten scenarios revealed substantial variation in participants' performance. The overall error rate varied widely, with some questions demonstrating strong comprehension and others exposing persistent challenges. The highest accuracy rates were observed in Q4 (89%), Q5 (85%), and Q7 (85%), indicating that participants were able to interpret the LGPD principles underlying these situations correctly. Conversely, Q8 (8%), Q9 (23%), and Q10 (50%) registered the lowest accuracy levels, suggesting that the legal nuances or contextual details in these scenarios were more complex or ambiguous. Particularly, Q8 presented the most critical gap, with only 2 correct answers out of 25, which may point to misconceptions about the proper handling of sensitive or exceptional data-processing cases. Overall, these results highlight that while most participants achieved high accuracy in the majority of scenarios, specific topics still require deeper clarification to ensure consistent understanding of LGPD compliance in educational contexts.

5.6 Discussion

During preliminary testing, a significant challenge was ensuring that the ChatGPT API adhered strictly to predefined instructions. The system was designed to guide teachers on LGPD compliance by providing concise and objective responses limited to three sentences and concluding the interactions decisively. However, ChatGPT often failed to interpret these instructions correctly, particularly during the final stages of interactions, leading to prolonged dialogs and responses that contradicted the goal of brevity. Addressing these issues required frequent prompt adjustments and continuous monitoring to align the AI's behavior with the system's objectives.

Additionally, usability issues affected the user experience in the preliminary evaluation, especially on mobile devices. Despite utilizing the Bootstrap framework for responsive design, the initial interface presented challenges on smaller screens. Specifically, the fixed scenario descriptions at the top of the page overlapped with the chat interface, hindering readability and interaction. To mitigate this, the scenario descriptions were integrated within the chat interface, allowing both elements to scroll together and keeping the latest messages visible. Although this adjustment improved navigation on mobile devices, further testing and detailed feedback from teachers are necessary to assess its effectiveness across various platforms.

The results of the case study show consistent improvements compared to the first evaluation. In the *Perceived Ease of Use* dimension, both the mean and median reached high values, reflecting the concentration of responses in the upper range of the scale (above 3.7). For *Perceived Usefulness*, the median was equal to 4.00, indicating a high degree of consensus among participants on the benefits provided by the system. Finally, the *Perceived Intention to Use* dimension showed a mean of 4.05, with a relatively uniform distribution of responses, suggesting a positive intention to continue using the LGPD Guide. These results highlight the effectiveness of the improvements made and their positive impact on the

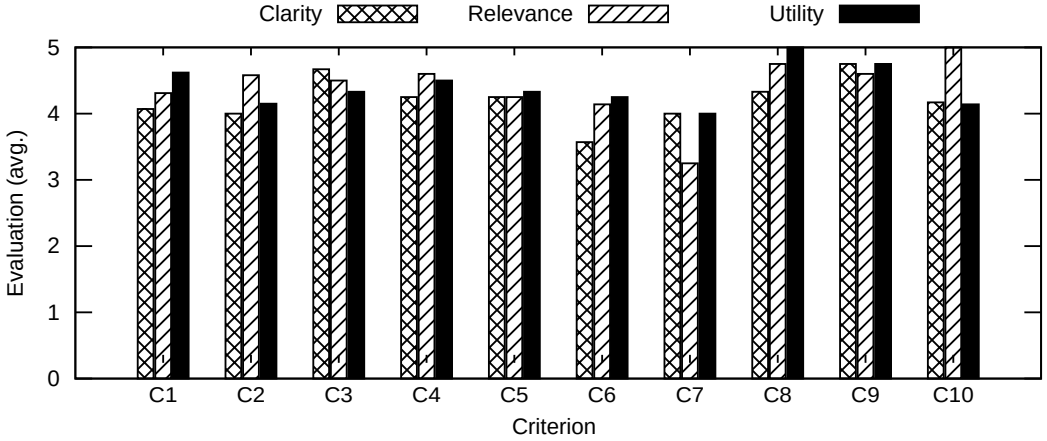

Figure 10. Average evaluations by criterion for each scenario.

educators' experience.

The results of the Mann–Whitney U tests revealed statistically significant differences between the two experiments across all three dimensions, considering a significance level of $p < 0.05$. Specifically, PEOU ($U = 1162.5, p = .002$), PU ($U = 243.5, p < .001$), and PIU ($U = 34.5, p = .042$) all showed significant differences in their distributions, indicating that participants in Experiment 2 reported higher perceptions across these constructs. Given that the data did not meet the assumptions of normality, the nonparametric Mann–Whitney test was used as an appropriate alternative to compare the independent samples. However, it is important to note that the unequal and relatively small sample size imposes limitations on the generalization of these findings.

In addition to the general evaluation conducted through the TAM, a detailed analysis of the interactions across different scenarios was performed using a star-rating system to assess the clarity, relevance, and usefulness of the responses generated by the LLM. Overall, the results revealed that usefulness was consistently well rated, often with averages above 4.0. Scenario 8, for instance, achieved the highest usefulness score (5.00), indicating that teachers considered its responses particularly applicable to their teaching practice.

Clarity was also positively evaluated, though some variation was observed—Scenario 6 recorded the lowest average (3.57), suggesting that its responses may have been more difficult to understand. Relevance, in turn, exhibited greater fluctuation between scenarios: Scenario 7 received the lowest rating (3.25), indicating weaker alignment with teachers' needs, whereas Scenario 10 attained the maximum score (5.00), demonstrating strong contextual adequacy.

These results indicate that, in general, the interactive guide was able to provide useful and clear responses; however, there is room for adjustments in certain scenarios, especially to ensure that the relevance of the responses is maintained in all cases. Additionally, future data collections could be organized to increase the number of evaluations in the less-responded scenarios and obtain an even more comprehensive view of the teachers' experiences.

5.6.1 Costs

Figure 11 shows the daily usage of the ChatGPT API throughout the month of November. A significant variation in activity is observed throughout the month, with two notable peaks around November 15 and 22, suggesting moments of higher teacher engagement. These peaks may be associated with specific actions, such as sending invitations or events that encouraged the use of the system. In contrast, there are periods of low activity, especially towards the end of the month, such as after November 27, possibly reflecting a decrease in interest or the sending of new invitations. Overall, the graph demonstrates a heterogeneous distribution in API usage, highlighting the influence of external factors on participant behavior during tests.

Analysis of API usage costs reinforces the economic feasibility of the experiment. The total cost recorded in November was just **$0.21**, reflecting the efficiency of the API over the period. This low cost is a direct result of the choice of the **gpt-3.5-turbo-0125** model, known for its efficiency and accessibility, charging **$0.0015 per 1,000 input tokens** and **$0.002 per 1,000 output tokens**. Considering the average of **12.16 interactions per session**, and assuming that each interaction uses about 50 to 100 tokens for combined input and output, the cost per session was quite low, reinforcing the model's suitability for this type of intensive-use experiment.

Furthermore, even on days with higher activity, such as November 15 and 22, the daily cost remained very low, demonstrating that the increase in the volume of interactions did not significantly impact the project's budget. Thus, the choice of the **gpt-3.5-turbo-0125** model was strategic in allowing a large number of interactions within a controlled cost, allowing for sufficient data collection for analysis without compromising available financial resources.

6 Conclusion

The LGPD represents a fundamental regulatory milestone for data security and privacy in Brazil, posing challenges for various educational institutions. Compliance with its guidelines requires not only a detailed understanding of the legislation,

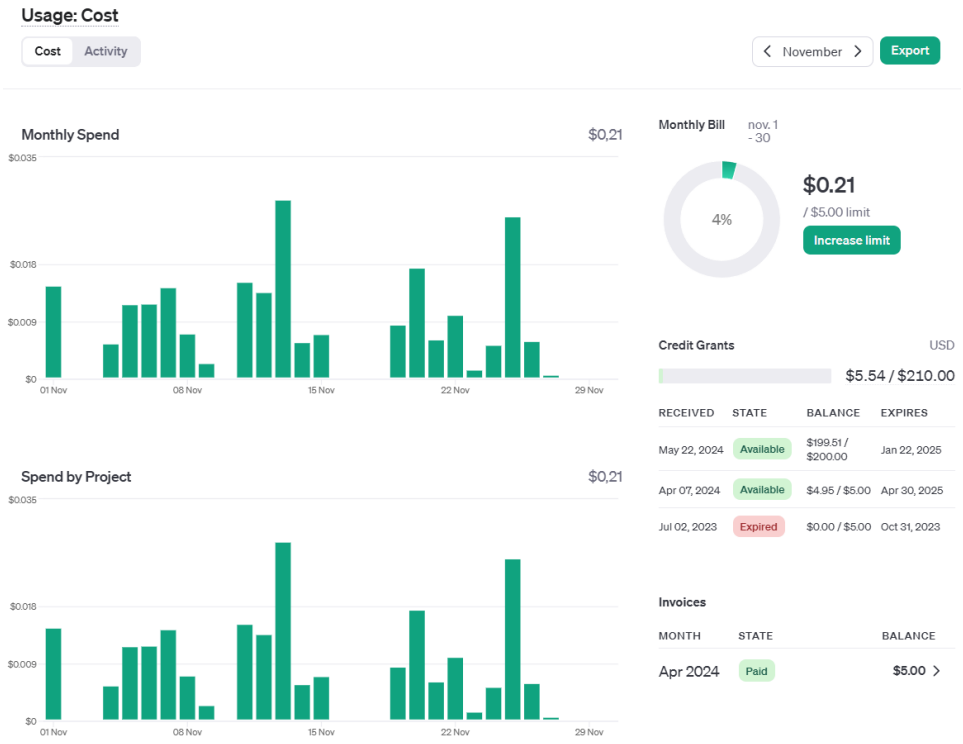

Figure 11. Chart of ChatGPT API usage throughout November.

but also the adoption of practices and tools that ensure conformity in the processing of personal data. In this context, teachers in educational institutions face difficulties in implementing the requirements of the LGPD, either due to a lack of specific knowledge about the regulation or the scarcity of accessible resources that facilitate this process. In response to this scenario, this study proposed the development of a data literacy approach through an interactive guide to help teachers understand the LGPD and adapt their daily tasks, using LLM to provide support tailored to the needs of the educational context.

In summary, the results achieved demonstrated that the development of the approach, combined with the application of evaluation methods such as TAM, allowed the attainment of the general objective of creating a tool to support teachers in understanding and adopting good practices related to the LGPD guidelines. The specific objectives—identifying requirements, implementing the system, and analyzing its acceptance—were met successfully. Through a preliminary evaluation and a case study, the effectiveness of the improvements made was evidenced.

Therefore, the results confirm that an interactive approach based on practical scenarios, supported by an LLM, can facilitate the understanding and application of the LGPD, promoting the dissemination of knowledge about legal compliance in teaching activities.

6.1 Main Contributions

The preliminary evaluation and the case study provided significant contributions to understanding how interactive technologies, such as the AI-based Interactive Guide, can be applied to teaching and compliance with the LGPD in the academic context. The main contributions of this study include:

1. **Evaluation of an Innovative Solution for LGPD Literacy:** The study evaluated the Interactive Guide as an effective tool for disseminating knowledge about the LGPD among teachers. The use of TAM revealed a considerable evolution in perceptions of Perceived Ease of Use, Perceived Usefulness, and Perceived Intention to Use, indicating that the system became more accessible and useful as usability and design adjustments were implemented.
2. **Impact of Iterative Improvements on the System:** The comparison between the preliminary evaluation and the case study demonstrated that iterative improvements, such as adjustments to the mobile interface and feedback collection strategies, resulted in a more positive user experience. These data suggest that a continuous development approach can significantly increase the acceptance and effectiveness of technological tools in teaching.
3. **Application of AI in the Educational and Regulatory Context:** This study contributes to an understanding of the use of language models such as ChatGPT in educational and regulatory contexts, highlighting the potential of this technology to resolve complex doubts and support teachers in applying the LGPD. The research also revealed critical challenges, such as the need to improve interaction control and prompt design – areas that require attention in future investigations.
4. **Identification of Critical Challenges for Implementing Educational Technologies:** In addition to positive results, the preliminary evaluation and case study iden-

Table 6. Number of evaluations and average by criterion for each scenario.

Scenario	Criterion	No. of Evaluations	Average
C1	Clarity	14	4.07
	Relevance	13	4.31
	Usefulness	13	4.62
C2	Clarity	11	4.00
	Relevance	12	4.58
	Usefulness	13	4.15
C3	Clarity	3	4.67
	Relevance	4	4.50
	Usefulness	6	4.33
C4	Clarity	4	4.25
	Relevance	5	4.60
	Usefulness	2	4.50
C5	Clarity	4	4.25
	Relevance	4	4.25
	Usefulness	3	4.33
C6	Clarity	7	3.57
	Relevance	7	4.14
	Usefulness	8	4.25
C7	Clarity	5	4.00
	Relevance	4	3.25
	Usefulness	3	4.00
C8	Clarity	3	4.33
	Relevance	4	4.75
	Usefulness	3	5.00
C9	Clarity	4	4.75
	Relevance	5	4.60
	Usefulness	4	4.75
C10	Clarity	6	4.17
	Relevance	5	5.00
	Usefulness	7	4.14

tified challenges related to the usability and responsive design of the system, which were crucial for the guide's acceptance. These findings indicate that when integrating new technologies into teaching, it is essential to focus on user experience and ensure that the platform is intuitive and accessible.

These contributions enhance the understanding of the use of interactive systems in teaching complex regulations such as the LGPD and offer valuable insights for future research and development practices in educational technologies.

6.2 Limitations and Future Work

Despite promising results, some limitations were identified that can be addressed in future work. Firstly, issues related to controlling interactions with the OpenAI API, using the gpt-3.5-turbo-0125 model, which in some cases did not follow specific instructions to end conversations, indicate the need to explore other AI models/tools or improve prompt engineering for greater predictability in responses. It is also worth noting that viable no-cost and even offline LLM options are emerging, which could broaden adoption in educational settings by reducing dependency on commercial APIs.

Additionally, the creation of scenarios is still performed manually and evaluated by experts. In the future, the system can be expanded to serve other professional categories beyond teachers, such as managers and IT professionals, adapting the LGPD Guide content to audiences with specific needs. Another possibility is to integrate more interactive resources, such as quizzes or explanatory videos, to increase engagement and information retention.

A promising idea would be to integrate the LGPD Guide with academic tools already used in educational institutions, such as modules in academic portals, plugins for Learning Management Systems (LMS), or data and privacy management systems. This integration could facilitate the adoption of the guide, ensuring that good data protection practices are incorporated into the daily routines of educators and school managers.

It would also be relevant to conduct tests in a larger number of institutions and with a more diverse sample of teachers, aiming to validate the results in varied contexts. Furthermore, the format for collecting user evaluations, both of the guide in general and of interactions in each scenario, can be evolved to allow for more qualitative and detailed feedback, not just based on a 1 to 5 scale. This would provide a better interpretation of the evaluations through user' justifications and suggestions.

Finally, future studies can investigate the impact of the Interactive Guide on teachers' behavior changes regarding LGPD compliance, using long-term metrics such as the reduction in the number of data breach incidents after adopting the guide.

In conclusion, the data literacy approach based on an interactive LGPD guide using LLM represents a significant advancement in the dissemination of knowledge about the legislation, especially in the academic context. The tool proved to be effective both in expanding educators' understanding of the LGPD and in the practical application of its principles in daily activities. Furthermore, the results highlight the great potential of interactive educational technologies, supported by AI, to foster legal compliance and innovate in teaching complex regulations.

Although promising results were achieved, this work also revealed important challenges, such as the need to improve interaction design and expand system accessibility. These points, along with the new directions indicated for future work, suggest that there is room to deepen the research and explore applications of the Guide in other contexts and audiences. Thus, by addressing gaps in teaching and practicing the LGPD, this study contributes to advancing discussions on the role of technology in education and transforming institutional practices.

Declarations

Authors' Contributions

César Murilo da Silva Junior: Conceptualization, Data curation, Investigation, Formal analysis, Writing – original draft. Silvio E. Quincozes and Rafael D. Araújo: Conceptualization, Supervision, Project administration, Writing – review & editing. Juliana Saraiva: Writing – review & editing, Validation. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors gratefully acknowledge the support of the PPGCO/FACOM at UFU for enabling this work, as well as the voluntary contributions of the faculty members who responded to the survey.

Availability of data and materials

The source code is available at `https://github.com/cesarmsj/LGPD-guide-using-LLM-for-teaching-activities`. The dataset generated and/or analyzed during the current study will be available upon request.

References

Amatriain, X. (2024). Prompt design and engineering: Introduction and advanced methods. *arXiv preprint arXiv:2401.14423v4*. Last revised on 5 May 2024 (version v4). Accessed on: 04 August 2025. DOI: 10.48550/arxiv.2401.14423.

Angular (2024). Angular: The modern web developer's platform. Available at: `https://angular.io`. Accessed on: 12 out. 2024.

ANPD (2022). Guia Orientativo sobre Tratamento de Dados Pessoais para Fins Acadêmicos e de Pesquisa. Autoridade Nacional de Proteção de Dados. Available at:`https://www.gov.br/anpd/pt-br/centrais-de-conteudo/materiais-educativos-e-publicacoes/web-guia-anpd-tratamento-de-dados-para-fins-academicos.pdf` Accessed on: 24 mar. 2025.

ANPD (2023). Estudo Técnico: A LGPD e o Tratamento de Dados Pessoais para Fins Acadêmicos e para a Realização de Estudos por Órgão de Pesquisa. Autoridade Nacional de Proteção de Dados. Available at:`https://www.gov.br/anpd/pt-br/centrais-de-conteudo/documentos-tecnicos-orientativos/estudo-tecnico-a-lgpd-e-o-tratamento-de-dados-pessoais-para-fins-academicos-e-para-a-realizacao-de-estudos-por-orgao-de-pesquisa.pdf/view` Accessed on: 24 mar. 2025.

Barile da Silveira, D., Oliveira, M. d. G. M. D. d., and Mozaner, V. C. (2024). Os impactos da regulação sobre privacidade e proteção de dados na segurança da informação: um estudo à luz da gdpr e da lgpd. *Revista de Direito Administrativo e Infraestrutura (RDAI)*, 8(30):49–65. DOI: 10.48143/rdai.30.silveira.

Bioni, B. R. (2021). *Proteção de dados pessoais: a função e os limites do consentimento*. Forense, Rio de Janeiro, 3 edition. Available at:`https://brunobioni.com.br/livros/protecao-de-dados/`. Accessed on: 24 mar. 2025.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R. B., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., *et al*. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258v3*. DOI: Last revised on 12 July 2022 (version v3).10.48550/arxiv.2108.07258 Accessed on: 17 October 2024.

Brasil (2014). Lei nº 12.965, de 23 de abril de 2014 – Marco Civil da Internet. *Diário Oficial da República Federativa do Brasil*. Available at: `https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm`. Accessed on: August 2025.

Brasil (2018). Lei Geral de Proteção de Dados Pessoais - Artigos 7º e 11º. Available at: `https://www.jusbrasil.com.br/topicos/200399417/artigo-7-da-lei-n-13709-de-14-de-agosto-de-2018`. Accessed on: 24 mar. 2025.

Brasil (2018). Lei nº 13.709, de 14 de agosto de 2018 – Lei Geral de Proteção de Dados Pessoais (LGPD). *Diário Oficial da República Federativa do Brasil*. Available at: `http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm`. Accessed on: August 2025.

Brasil (2022). Emenda constitucional nº 115, de 10 de fevereiro de 2022. *Diário Oficial da União*. Available at: `https://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc115.htm`. Accessed on: August 2025.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are few-shot learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA. Curran Associates Inc.. DOI: 10.48550/arxiv.2005.14165.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Tandon, N., and Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712v5*. Last revised on 13 April 2023 (version v5). Accessed on: 4 August 2025. DOI: 10.48550/arxiv.2303.12712.

Costa Filho, B. A. d., Pires, P. J., and Hernandez, J. M. d. C. (2007). Modelo Technology Acceptance Model - TAM aplicado aos Automated Teller Machines - ATM'S. *Revista de Administração e Inovação*, 4(1):40–56. Available at:`https://revistas.usp.br/rai/article/view/79071`.

Coursera (2023). Coursera. Available at: `https://coursera.org/`. Accessed on: 17 October 2024.

da Silva, L. M. M. and Takeshita, L. M. A. (2023). Lei geral de proteção de dados pessoais e a preservação da intimidade do professor. *Prisma Juridico, v22, n. 1*, pages 23–39. DOI: 10.5585/2023.19473.

DAMA International (2023). Data governance framework. Accessed on: 17 October 2024. DOI: 10.1002/9781119851462.ch2.

Davis, F. D. (1985). *A Technology Acceptance Model for Empirically Testing New End-user Information Systems: Theory and Results*. PhD thesis, Massachusetts Institute of Technology, Sloan School of Management. Available

at:https://dspace.mit.edu/handle/1721.1/15192.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340. DOI: 10.2307/249008.

de Almeida, A. O., Junior, A. A., Canato, R. L. C., Albardeiro, S. T., and Marques, V. C. (2020). Ética, segurança e privacidade na educação à distância durante a pandemia no Brasil. *Revista InovaEduc*, (7):1–28. Book.

de Almeida, W. G., de Souza Araújo, N. V., and Oprime, P. C. (2023). Uma proposta de metodologia para adequação à lgpd em uma instituição federal de ensino superior. In *Anais da XII Escola Regional de Informática de Mato Grosso*, pages 79–86. SBC. DOI: 10.5753/eri-mt.2023.236499.

Doneda, D. (2019). *Da Privacidade à Proteção de Dados Pessoais*. Revista dos Tribunais (edição Renovar). Book.

dos Santos, S. R. R. and da Silva Pereira, S. M. (2023). Lei Geral de Proteção de Dados Pessoais (LGPD): Efetividade das ações desenvolvidas pela Universidade Federal de Sergipe (UFS). *Revista Expectativa*, 22(4):49–74. DOI: 10.48075/revex.v22i4.30653.

EDPB (2023). GDPR compliance toolkit. Available at:https://www.dpa.gr/en/byDesign_onlinetoolkit_en. Accessed on: 17 October 2024.

edX (2023). E-learning platform. Available at:https://edx.org/. Accessed on: 17 October 2024.

European Union (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Available at:https://eur-lex.europa.eu/eli/reg/2016/679/oj. Accessed on: August 2025.

Feitosa Jr., A. (2020). Falha em site da anhembi morumbi permite acessar dados pessoais de alunos. Available at:https://gizmodo.uol.com.br/falha-anhembi-morumbi-dados-pessoais-alunos Accessed on: October 2023.

Firebase (2024). Firebase documentation. Available at: https://firebase.google.com/docs. Accessed on: 12 out. 2024.

Git Development Community (2024). Git documentation. Available at: https://git-scm.com/doc. Accessed on: 12 October 2024.

GitHub, I. (2024). Github documentation. Available at: https://docs.github.com/en/. Accessed on: 12 October 2024.

Gomes, F. V., Filho, M. C. C., and Luccas, V. N. (2023). Proteção de dados e instituições de ensino: o que fazer com dados de alunos? *Revista Brasileira de Políticas Públicas . abr2023, Vol. 13 Issue 1*, pages p402–420. 19p.. DOI: 10.5102/rbpp.v13i1.7996.

Hoel, T. and Chen, W. (2018). Privacy and data protection in learning analytics should be motivated by an educational maxim—towards a proposal. *Research and Practice in Technology Enhanced Learning*, 13(1):20. DOI: 10.1186/s41039-018-0086-8.

Jekabsone, I. (2023). Selected legal issues in online adult education: Compliance of online learning and teaching process with gdpr. *TalTech Journal of European Studies*, 13(2):46–62. DOI: 10.2478/bjes-2023-0015.

Jesus, D. C. d. (2022). Proposta de um projeto de conformidade a partir das práticas da ISO 27701 para implementação de um programa compliance de proteção de dados à luz da LGPD na universidade de rio verde. Master's thesis, Universidade do Vale do Rio dos Sinos. Defense on November 23, 2022. Open access via Unisinos Digital Repository, BDTD – Brazilian Digital Library of Theses and Dissertations Available at:http://www.repositorio.jesuita.org.br/handle/UNISINOS/12162.

Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260. DOI: 10.1126/science.aaa8415.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22:5–55. Available at:https://psycnet.apa.org/record/1933-01885-001.

Marín, V. I., Carpenter, J. P., Tur, G., and Williamson-Leadley, S. (2023). Social media and data privacy in education: an international comparative study of perceptions among pre-service teachers. *Journal of Computers in Education*, 10:769–795. DOI: 10.1007/s40692-022-00243-x.

Marques, A. F. (2021). A implantação da lei geral de proteção de dados na universidade federal do rio grande do sul: uma análise a partir da noção de regime de informação e seus componentes. Master's thesis, Universidade Federal do Rio Grande do Sul. Available at:https://lume.ufrgs.br/handle/10183/249811Published in Portuguese.

Martins, M. E., Aguiar, Y. P. C., and Saraiva, J. (2025). Assessment of competences for lgpd dpo through anpd standard and information systems curriculum. In *Anais do XXI Simpósio Brasileiro de Sistemas de Informação (SBSI)*, pages 565–574. DOI: 10.5753/sbsi.2025.246585.

Novais, A. F. S., de Oliveira, E. C., da Silva, F. B. S., Júnior, H. G. M., Demuner, J. A., da Silva, K. F. T., da Silva, M. L. L., Holanda, M. G., and de Souza, M. M. (2024). Promovendo segurança online no ambiente educacional moderno. *REVISTA FOCO*, 17(1):e4113–e4113. DOI: 10.54751/revistafoco.v17n1-049.

Oliveira da Silva, K. and Costa Sarkis, L. (2024). Conformidade da lgpd em instituições públicas de ensino superior no brasil: as perspectivas de profissionais de tic. *Cadernos do IME - Série Informática*, 49:9–26. DOI: 10.12957/cadinf.2024.80214.

OneTrust (2023). Soluções onetrust para a LGPD. Available at: https://www.onetrust.com/br/lgpd/. Accessed on: 17 October 2024.

OpenAI (2024). ChatGPT. Available at: https://www.openai.com/chatgpt. Accessed on: 12 out. 2024.

Papamitsiou, Z., Filippakis, M. E., Poulou, M., Sampson, D., Ifenthaler, D., and Giannakos, M. N. (2021). Towards an educational data literacy framework: enhancing the profiles of instructional designers and e-tutors of online and blended courses with new competences. *Smart Learning Environments*, 8:18. DOI: 10.1186/s40561-021-00163-w.

Pinheiro, P. P. (2023). *Proteção de Dados Pessoais. Comentários à Lei N. 13.709/2018 (LGPD)*. Saraiva, 4 edition. Book.

Pinto, A. L. d. S., Torres, E. F., Moura, J. M. d., Sousa,

E. d. S., Pinto, L. A., and Nóbrega, C. V. d. (2019). Avaliação da aceitação das ferramentas tecnológicas no ambiente do trabalho docente. *Revista Gestão Universitária na América Latina - GUAL*, 12(2):118–138. DOI: 10.5007/1983-4535.2019v12n2p118.

PROPLAD/UFU (2025). Anuário 2024 da Universidade Federal de Uberlândia. Available at: `https://proplad.ufu.br/central-de-conteudos/documentos/2025/05/anuario-2024`. Accessed on: 27 September 2025.

Reis, S. R. F., Reis, T. S. M., de Almeida, G. L., Júnior, T. A. F., dos Santos Farias, A. S., Matos, H. C. S., Matos, L. S., Feitoza, M. L., Fernandes, A. B., and Fukuoka, D. M. L. (2024). Desafios da LGPD quanto à privacidade em ambientes educacionais: um mapeamento sistemático. *Revista de Gestão e Secretariado*, 15(3):e3292–e3292. DOI: 10.7769/gesec.v15i3.3292.

Rohr, A. (2021). Universidade do RS se desculpa por vazar dados de 23 mil alunos. Available at: `https://l1nk.dev/1r8wn`. Accessed on: October 2023.

Rojas, M. A. T. (2020). Avaliação da adequação do Instituto Federal de Santa Catarina à lei geral de proteção de dados pessoais. Trabalho de conclusão de curso, Centro de Referência em Formação e Educação a Distância, Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina, Santa Catarina. Available at:`https://zenodo.org/records/4445239`.

Saraiva, J., de Souza, C., and Soares, S. C. B. (2025). MMAI-LGPD: A maturity model for governance and data compliance in information systems institutions. In *Anais do XXI Simpósio Brasileiro de Sistemas de Informação (SBSI)*. DOI: 10.5753/sbsi.2025.246644.

Saraiva, J. d. A. G., Araújo, J., and Soares, S. C. B. (2024). Ensino da adequação à lgpd no desenvolvimento de software através da aprendizagem ativa e centrada no discente. In *Anais do IV Simpósio Brasileiro de Educação em Computação (EduComp)*, pages 204–213. DOI: 10.5753/educomp.2024.237528.

Saraiva, J. d. A. G. and Soares, S. C. B. (2023). Privacy and security documents for agile software engineering: An experiment of LGPD inventory adoption. In *ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2023)*, pages 1–9. DOI: 10.1109/ESEM56168.2023.10304806.

Silva Junior, C. F. d. (2020). Cidadania digital na prevenção de ciberbullying. Dissertação (mestrado em comunicação), Universidade Municipal de São Caetano do Sul (USCS), São Caetano do Sul, SP, Brasil. Programa de Pós-Graduação em Comunicação, Mestrado Profissional em Inovação na Comunicação de Interesse Público Available at:`https://repositorio.uscs.edu.br/handle/123456789/921`.

Souza, J. G. S., Arima, C. H., and Belda, F. R. (2020). Análise de tratamento da segurança da informação de uma instituição de ensino público federal. *RIAEE - Revista Ibero-Americana de Estudos em Educação, Araraquara*, 15(3):1309–1321. DOI: 10.21723/riaee.v15i3.13584.

Souza, J. G. S., Arima, C. H., and Belda, F. R. (2022). Análise de aplicação da lgpd numa instituição pública de ensino: Um estudo de caso. *RIAEE - Revista Ibero-Americana de Estudos em Educação, Araraquara*, 17(3):1856–1872. DOI: 10.21723/riaee.v17i3.16789.

Stelzer, J., das Neves Gonçalves, E., Baptista, R. R. F., Vaz, R. M. P., Wieira, K., and de Medeiros Fidelis, M. (2019). A lei geral de proteção de dados pessoais e os desafios das instituições de ensino superior para a adequação. In *XIX Colóquio Internacional de Gestão Universitária*, Florianópolis. UFSC. Available at:`https://repositorio.ufsc.br/handle/123456789/201939`.

Teodoro, J., Oliveira, L. S., Junior, J. F. S., Senes, R. J., and Augusto, T. C. (2023). Um modelo canvas do processo de adaptação à lei geral de proteção de dados: o caso da universidade do estado de santa catarina (udesc). *Revista Brasileira de Biblioteconomia e Documentação*, 19:1–28. DOI: 10.58876/rbbd.2023.1911869.

Torres-Hernández, N. and Gallego-Arrufat, M.-J. (2023). Pre-service teachers' perceptions of data protection in primary education. *Contemporary Educational Technology*, 15(1):ep399. DOI: 10.30935/cedtech/12658.

TrustArc (2023). Trustarc: Privacy management software. Available at:`https://trustarc.com/`. Accessed on: 17 out. 2024.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 6000–6010. Curran Associates Inc.. DOI: 10.65215/mdcm8z23.

Venkatesh, V. and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2):186–204. DOI: 10.1287/mnsc.46.2.186.11926.

Wendling, G. S., Vanini, L., Fortes, M. C., Wiest, R., and Rossetto, A. G. d. M. (2023). Diagnóstico do nível de maturidade da aplicação da lgpd nas escolas de educação infantil da rede municipal de educação de passo fundo. *CONTRIBUCIONES A LAS CIENCIAS SOCIALES*, 16(8):11359–11376. DOI: 10.55905/revconv.16n.8-135.

Zappelini, T. D., Schippers, L.-M., and da Silva, A. P. (2023). Personal data of adolescents in higher education: The LGPD and the use of digital media for teaching activities and remote testing. *Revista Direito GV*, 19. DOI: 10.1590/2317-6172202335.