Investigating Privacy by Design in Software Development: Insights from Brazilian Federal Higher Education Institutions

Fernando Elias de Oliveira[1], **Edna Dias Canedo**[1]

[1]University of Brasília (UnB),
Professional Postgraduate Program in Privacy and Information Security (PPEE)
e-mail: fernandoeliasti@gmail.com, ednacanedo@unb.br

Abstract. ***Research Context:*** *The increasing digitalization of services in Brazilian Federal Higher Education Institutions (IFES) intensifies the processing of sensitive personal data, raising urgent concerns regarding compliance with the General Data Protection Law (LGPD) and the systematic adoption of Privacy by Design (PbD). Although privacy is recognized as a constitutional right and a central element of data governance, audits and prior studies reveal a lack of consolidated methods to assess and improve privacy maturity in software development within IFES.* ***Practical Problem:*** *Despite the existence of PbD principles and international frameworks, organizations still face difficulties in translating abstract legal requirements into concrete software engineering practices. This gap is particularly critical in IFES, where academic management systems store highly sensitive information and where deficiencies in privacy and security have been repeatedly identified by oversight bodies.* ***Proposed Solution:*** *This study investigates how IT professionals in IFES perceive, understand, and apply PbD principles in their daily work. Through a mixed-method survey, we diagnose the maturity of privacy practices across the software lifecycle, identifying adoption levels, barriers, and enablers of PbD implementation.* ***Related IS Theory:*** *The research is grounded in socio-technical perspectives of Information Systems, drawing on Privacy Engineering, Privacy Requirements Engineering, and adoption models such as the Unified Theory of Acceptance and Use of Technology (UTAUT). These theoretical lenses highlight the interplay between individual perceptions, organizational culture, and technical practices in shaping PbD maturity.* ***Research Method:*** *A survey instrument with 55 questions, including Likert-scale and open-ended items, was distributed to IT professionals in IFES. Responses from 58 participants across 15 Brazilian states were analyzed using descriptive statistics, correlation analysis, and qualitative coding based on Grounded Theory.* ***Summary of Results:*** *The findings reveal strong awareness of privacy as a fundamental right and recognition of shared responsibility, but also significant gaps in regulatory knowledge, structured training, and systematic adoption of advanced privacy strategies. While encryption and minimization are common, strategies such as decentralization, sovereignty, and proxies remain underutilized. Moreover, organizational structures are fragmented, tools are rarely adopted, and practices tend to be reactive rather than proactive.* ***Contributions and Impact to IS area:*** *This study provides empirical evidence of how PbD is understood and practiced in IFES, exposing maturity gaps and socio-technical barriers. It contributes a diagnostic perspective that informs both academia and practice, supporting the design of frameworks and*

instruments to foster PbD adoption in public-sector information systems, where sensitivity of data and regulatory pressures are particularly high.
Keywords: *Privacy by Design; LGPD; Privacy Engineering; Software Development Lifecycle; Brazilian Federal Higher Education Institutions.*

1. Introduction

The preservation of privacy is currently at the center of global agendas and organizational processes in both the public and private sectors [Andrade et al. 2024]. Article 12 of the Universal Declaration of Human Rights directs governments to review procedures, practices, and legislation related to communication surveillance, interception, and the collection of personal data [United Nations 2025]. Scholars, activists, and policymakers consistently argue that individuals' rights must be safeguarded in both the physical and digital worlds, with privacy being a fundamental right [Chander and Land 2014].

The rapid advancement of the internet has led to an increasing collection and detailed profiling of individuals' data [Spósito et al. 2025]. This development highlights the urgent need for rigorous privacy-preserving techniques in information technology processes, enabling society to enjoy the benefits of the internet in a secure manner [Spósito et al. 2025, Chander and Land 2014]. High-profile cases such as the Cambridge Analytica–Facebook scandal have demonstrated how personal data can be exploited for commercial, political, and strategic purposes by those who control it [Confessore 2018].

More recently, in 2024, a massive breach known as RockYou2024 exposed 10 billion user credentials online, ranking among the largest data leaks to date [Rodrigues et al. 2025, Alexey Andreev 2025]. Social networks and technology platforms have been recurrent victims of cyberattacks and data breaches: in 2019, over 533 million Facebook users were affected, while Yahoo suffered the exposure of more than 3 billion accounts, remaining unaware of the breach for years [Alexey Andreev 2025]. In Brazil, over 223 million citizens had sensitive data—such as addresses, taxpayer IDs, social benefits, credit scores, and income tax records—leaked online in 2021, with evidence suggesting that the source was the SERASA platform [André Luiz Dias Gonçalves 2025]. Furthermore, in 2024 the number of data breaches in the Brazilian government was over 20 times higher than in 2020, exceeding 9,000 reported incidents [Rodrigues et al. 2024]. Although these platforms collected and shared personal data to provide services, users were ultimately left in vulnerable situations due to repeated exposures [Spiekermann 2012].

Safeguarding privacy in the context of organizational information systems is a complex issue that involves multiple domains of knowledge and stakeholders [Matos et al. 2025]. The development of new technologies must, therefore, consider mechanisms to prevent or mitigate risks that compromise the fundamental right to data protection, as guaranteed by the Brazilian Constitution [Alves and Neves 2021]. Cavoukian [Cavoukian 2012], pioneer of the concept of Privacy by Design (PbD), argues that embedding privacy into system design is not merely desirable but necessary to address the emerging threats to privacy in a rapidly evolving technological environment.

PbD is an approach that advocates for the integration of data protection measures into systems from their inception, rather than as post-hoc adjustments. It is grounded in seven core principles: proactive not reactive; privacy as the default setting; privacy em-

bedded into design; full functionality (positive-sum, not zero-sum); end-to-end security; visibility and transparency; and respect for user privacy [de Chaves and Benitti 2023].

In Brazil, the General Data Protection Law (LGPD) established strict guidelines for the processing of personal data [Presidência da República do Brasil 2018]. In this context, PbD provides a methodology for embedding privacy principles into the development of information systems and services. PbD principles—such as data minimization, purpose limitation, and security by design—are central to achieving compliance with the LGPD [Rocha et al. 2023].

However, audits conducted by the Federal Court of Accounts (TCU) in recent years reveal that both Federal Higher Education Institutions (IFES) and oversight bodies lack consolidated methods to evaluate and monitor privacy maturity [Tribunal de Contas da União 2025]. This gap raises the following research question: What is the level of maturity in the implementation of Privacy by Design principles within the information systems of Brazilian Federal Higher Education Institutions? To address this question, the present study aims to evaluate the adoption of PbD practices in the software processes of academic management systems at IFES in Brazil.

The increasing digitalization of services in IFES has intensified the need for robust data protection practices aligned with the LGPD, as reinforced by Ordinance SGD/MGI No. 852 of March 28, 2023 [Secretaria de Governo Digital 2023]. Nevertheless, studies indicate that many organizations continue to struggle with LGPD compliance, often due to the absence of consolidated methodologies or to technical and organizational challenges [Rocha et al. 2023]. This study seeks to contribute to advancing privacy maturity in IFES by proposing a framework to support the evaluation and improvement of privacy practices in software processes.

The absence of systematic approaches that integrate privacy into software development processes within IFES may compromise the protection of personal data—such as academic records, financial data, medical information, and research participation records. Without the incorporation of principles such as minimization, access control, purpose limitation, and security by design from the earliest stages of the software lifecycle, these data may be exposed or misused [Menegazzi and Silva 2023]. Such gaps not only pose legal risks under the LGPD but also threaten the integrity of core institutional activities, including education, research, and administration. Embedding privacy engineering into development processes is, therefore, a strategic requirement to ensure regulatory compliance, mitigate risks, and strengthen institutional resilience to privacy incidents [Bu et al. 2020].

TCU audits report that over 70% of IFES exhibit deficiencies in information security and data protection, exposing sensitive academic and socioeconomic data of students and staff [Tribunal de Contas da União 2025]. Previous breaches have already caused reputational harm and triggered the risk of sanctions under the LGPD, which establishes fines of up to R$ 50 million per violation [Presidência da República do Brasil 2018]. In this scenario, integrating privacy into the software development lifecycle is important for mitigating legal risks, protecting the trust of the academic community, and safeguarding the continuity of institutional activities.

2. Background and Related Work

Within the field of privacy engineering, several efforts have been undertaken to systematize the incorporation of data protection into the software development lifecycle [Sangaroonsilp et al. 2023]. Privacy Engineering (PE) methodologies have been proposed to support this integration, encompassing techniques such as Privacy Impact and Risk Assessments (PIRA), Model-based Development (MbD), and Privacy Requirements Engineering (PRE) [Al-Slais 2020]. Despite these initiatives, there remains a persistent gap in holistic and systematic methodologies capable of effectively translating privacy requirements into engineering activities, particularly in practical contexts.

Complementary to these approaches, the literature has explored the categorization and classification of privacy requirements. Examples include taxonomies based on LGPD and ISO/IEC 29100, as well as taxonomies designed to mine and classify privacy requirements from issue reports [Ferrão et al. 2024]. A systematic mapping study on the application of Privacy by Design (PbD) in software engineering revealed that the field remains immature: most publications focus on requirements and architectures, while empirical validation of proposed approaches is still lacking [Morales-Trujillo et al. 2018]. Although these contributions advance the conceptualization of privacy requirements and provide useful taxonomies, the literature remains scarce in offering concrete instruments to diagnose and measure how PbD principles are incorporated into academic or administrative systems. This gap constitutes the central focus of the present study.

The successful implementation of privacy in software development extends beyond technical and methodological aspects, being significantly influenced by behavioral and organizational factors that shape the adoption of new practices and change management [Bu et al. 2020]. Research has investigated information systems professionals' perceptions of privacy implementation through technology acceptance models, such as the Unified Theory of Acceptance and Use of Technology (UTAUT) [Bu et al. 2021]. Studies indicate that performance expectancy, perceived effort, social influence, and facilitating conditions strongly shape the intention to adopt PbD practices [Bu et al. 2020, Bu et al. 2021].

In addition, factors such as awareness of privacy, fear of sanctions for non-compliance, and perceived rewards for adoption act as motivators or barriers to the acceptance and effective use of privacy practices [Bu et al. 2020, Bu et al. 2021]. The perception that privacy implementation may increase workload or complexity—contrasted with a clear awareness of its benefits and legal implications—highlights the importance of approaches that take into account both developers' and organizations' perspectives [Bu et al. 2020, Bu et al. 2021]. Consequently, beyond theoretical guidelines and frameworks, the literature emphasizes the need to understand organizational and behavioral dynamics to overcome the challenges of integrating privacy, particularly in the context of academic and administrative systems, where instruments to diagnose such maturity from a PbD perspective remain scarce.

The effectiveness of privacy implementation in software development, therefore, does not reside solely in the application of technical guidelines but is intrinsically linked to the behavioral and organizational factors that shape how privacy is perceived and practiced [Ribak 2019]. Developers' mindsets, for example, play a critical role: many tend to conflate security with privacy, which not only limits their ability to embed adequate

protections but also shifts the responsibility for privacy to others [Hadar et al. 2018]. Such misunderstandings can lead to superficial implementation of data protection mechanisms, failing to address privacy nuances from the outset of the software lifecycle [Hadar et al. 2018]. Moreover, organizational culture and implicit norms within development environments significantly influence how privacy principles are interpreted and applied in practice [Ribak 2019].

Translating abstract privacy requirements into the day-to-day practices of developers is a complex process that involves mediating different development cultures, in which the very notion of privacy may be renegotiated and adapted [Ribak 2019]. The perceptions of IT professionals regarding privacy in software, and the extent to which these perceptions align with organizational expectations and needs, are critical to advancing privacy maturity [Canedo et al. 2020]. Nevertheless, the literature still lacks a deeper understanding of how individual developers' perceptions and organizational privacy culture interact to shape the maturity of PbD adoption—especially in the specific context of academic and administrative systems. Furthermore, there is a shortage of systematic instruments to diagnose and measure this incorporation of PbD principles in practice.

Recent studies further illustrate these gaps. Spósito et al. [Spósito et al. 2025] conducted a systematic literature review of 125 primary studies combined with a survey of 37 practitioners to identify techniques, methods, processes, frameworks, and tools addressing privacy requirements across the phases of requirements engineering. Their results revealed a strong emphasis on academic contexts, with approaches such as PriS, Secure Tropos, LINDDUN, STRAP, and SQUARE applied predominantly in the elicitation phase, while industrial adoption remains limited due to lack of awareness and practical validation. Rocha and Canedo [Rocha et al. 2023] extended this discussion by comparing major privacy laws (LGPD, GDPR, ADPPA, and the Australian Privacy Act) and frameworks (PbD and ISO/IEC 29100), using a systematic literature review, survey, and framework analysis. They identified 18 classes of challenges in ensuring compliance—such as lack of knowledge, ambiguity of laws, and absence of guides or standardized tools—highlighting the persistent difficulty of translating abstract legal principles into actionable technical practices. A third study [Matos et al. 2025] focused on organizational and behavioral dimensions of PbD adoption, emphasizing how developers' misconceptions between security and privacy, together with cultural norms embedded in organizations, lead to superficial or inconsistent integration of privacy principles. By analyzing these socio-technical dynamics, the authors argue that advancing PbD maturity requires not only technical frameworks but also instruments capable of capturing developers' perceptions and organizational culture. Andrade et al. [Andrade et al. 2024] addressed this methodological gap by mapping the seven PbD principles against 72 Privacy Patterns cataloged by UC Berkeley, using a structured consensus process with three experts. Their findings demonstrated strong correlations between Privacy Patterns and PbD principles—especially Respect for User Privacy, *Visibility and Transparency*, and *Privacy Embedded into Design*—and provided an empirically grounded instrument that helps bridge abstract principles with concrete software engineering practices. Finally, Andrade et al. [Andrade et al. 2023] conducted a multiple case study with ten practitioners across five organizations from different sectors, analyzing how personal data privacy is integrated into software development processes. Their results revealed that privacy concerns are typically addressed only in later stages of development, teams lack privacy specialists,

tools are rarely employed, and privacy is often conflated with security. These empirical insights underscore the immaturity of privacy practices in industry and the urgent need for systematic approaches to embed PbD into development processes.

In summary, prior research highlights five interrelated gaps: (i) the lack of systematic and empirically validated methods to embed privacy into engineering activities, (ii) the persistent difficulty of operationalizing legal requirements across different jurisdictions and aligning them with software development practices, (iii) the influence of behavioral and organizational factors on how privacy is implemented in practice, (iv) the need for concrete instruments to translate abstract PbD principles into actionable engineering practices, and (v) the evidence that organizations continue to treat privacy reactively, often confusing it with security and lacking dedicated specialists or tools. Addressing these gaps is important for developing diagnostic instruments that can measure and foster PbD maturity in the specific context of academic and administrative systems, which constitutes the primary contribution of this study.

3. Study Settings

This study aims to investigate the perceptions of Information Technology (IT) professionals from Brazilian Federal Higher Education Institutions (IFES) regarding the practices adopted to implement data privacy in software development processes. In particular, the research seeks to understand how software development professionals in these institutions conceptualize data privacy, how they adopt privacy-preserving practices in their daily activities, and which organizational and individual factors influence such adoption.

To achieve this objective, we designed and applied a survey composed of both closed- and open-ended questions, covering aspects of knowledge, attitudes, behaviors, strategies, and organizational context related to privacy in the software lifecycle. Based on the literature review and the survey design, we formulated the following Research Questions (RQs):

- RQ.1 How do IT professionals in Brazilian Federal Higher Education Institutions perceive and understand data privacy in the context of software development?
  This research question aims to capture participants' conceptual knowledge, awareness of risks, and attitudes toward privacy as a right and as a professional responsibility.
- RQ.2 What practices, techniques, and strategies are adopted by IT professionals in Brazilian Federal Higher Education Institutions to implement data privacy throughout the software development lifecycle?
  This research question investigates how professionals actually integrate privacy into daily activities, including identifying problems, proposing solutions, and applying privacy-enhancing strategies (e.g., encryption, minimization, anonymization).
- RQ.3 What organizational, technical, and individual factors influence the adoption and maturity of data privacy practices in software development at Brazilian Federal Higher Education Institutions?
  This research question seeks to identify barriers and enablers such as the existence of dedicated privacy teams, the availability of tools, management support, productivity concerns, and recognition or sanctions associated with privacy practices.

To address the research questions, we conducted a survey with 58 practitioners from different Federal Higher Education Institutions. The following sections present the study design in detail, including the characterization of the target audience, the structure of the questionnaire, the pilot test, the procedures for invitation and distribution, and the strategies adopted for data analysis.

3.1. Target Audience

The target audience of this study consists of Information Technology (IT) professionals working in Brazilian Federal Higher Education Institutions (IFES). Focusing specifically on IT practitioners is important, as these professionals are directly responsible for implementing techniques, processes, and methodologies that ensure user data privacy in the software systems developed by their teams.

To guarantee that participation was restricted to IT professionals from IFES, several measures were adopted. The invitation and introductory instructions of the questionnaire explicitly stated the intended audience. The survey was strategically distributed through institutional email addresses of IT professionals available on official IFES websites. Additionally, a screening question was included at the beginning of the survey to exclude participants who did not meet the established criteria. The survey questions themselves were also carefully designed with direct references to the target group. Furthermore, invitations were reinforced by sending direct messages to practitioners' profiles on social media platforms and, when available, through WhatsApp. Participants were also encouraged to share the survey with their colleagues, broadening the outreach within the target audience.

3.2. Survey Design

The survey was designed to capture the perceptions and practices of IT professionals in Brazilian Federal Higher Education Institutions regarding data privacy in software development. It was structured into six sections, covering both demographic information and thematic aspects of privacy awareness, attitudes, behaviors, and organizational practices. The instrument contained a total of 55 questions, distributed as follows:

1. Participant Profile (Q1–Q10): This section collected demographic and professional background information, including age, education level, role, seniority, type of employment, and professional experience with data privacy. These questions were all closed-ended, except for one open-ended item (Q10) allowing participants to briefly describe their prior experience with data privacy.
2. Privacy Awareness – Knowledge (Q11–Q21): This section focused on participants' conceptual understanding and knowledge of privacy. It included items about familiarity with privacy laws and standards (e.g., LGPD, ISO/IEC 29100, PbD), training, self-learning practices, and recognition of privacy-related risks. One open-ended question (Q13) asked participants to list five words that come to mind when thinking about privacy, while the remaining were closed-ended Likert scale items.
3. Privacy Awareness – Attitudes and Sentiments (Q21–Q27): This block measured professionals' personal attitudes and beliefs regarding privacy, such as distinguishing between security and privacy, concerns about monitoring, valuing informed

consent, or perceiving privacy as a fundamental right. Most items were closed-ended Likert scales, with one optional open-ended question for participants to justify strong disagreements.

4. Privacy Behaviors and Actions (Q28–Q37): This section investigated concrete actions performed by participants in their professional context, such as recognizing data retention limits, identifying privacy problems, proposing solutions, or advocating privacy in their teams. These items were primarily closed-ended Likert scale questions, complemented by an optional open-ended question to explain disagreement.
5. Privacy Strategies and Practices (Q38–Q43): This part examined the frequency, importance, and ease of use of privacy strategies and techniques (e.g., encryption, minimization, anonymization, risk management, code reviews). Questions were closed-ended matrix items with Likert-type scales for frequency, importance, and perceived ease of use. An additional open-ended question (Q40, Q42) allowed participants to suggest other strategies or justify their assessments.
6. Organizational and Individual Factors (Q44–Q55): This final section addressed organizational dynamics and adoption factors, such as the existence of dedicated privacy teams, use of tools, prioritization of privacy, productivity impacts, incentives, sanctions, and perceptions about Privacy by Design. Most were closed-ended questions (Likert scales and categorical options), with one open-ended item (Q45, Q48) asking respondents to describe their organization's privacy dynamics and identify major future challenges.

In total, the survey combined 46 closed-ended questions and 9 open-ended questions. The closed-ended items primarily used five-point Likert scales or multiple-choice options, ensuring comparability of responses, while the open-ended items enabled deeper qualitative insights into participants' perceptions, experiences, and contextual practices, as shown in Table 1. The full questionnaire, including all response options, as well as the anonymized responses of the participants, is openly available on Zenodo at `https://zenodo.org/records/17247244`.

Table 1. Survey questions, type, and associated Research Question (RQ)

ID	Question	Type	RQ
Q1	What is your age group?	Multiple choice	Profile
Q2	In which state is your institution located?	Multiple choice	Profile
Q3	What is your highest completed level of formal education?	Multiple choice	Profile
Q4	What is your work model (remote, hybrid, on-site)?	Multiple choice	Profile
Q5	Which role best describes your current activities in software lifecycle projects?	Multiple choice	Profile
Q6	Indicate the seniority level of your current position.	Multiple choice	Profile
Q7	How many years of experience do you have in IT/software development roles?	Multiple choice	Profile
Q8	What is your type of employment with the institution?	Multiple choice	Profile
Q9	Do you have or have you ever had professional experience related to data privacy?	Multiple choice	Profile
Q10	Briefly describe your experience with data privacy in software. If none, state that you have no experience.	Open-ended	Profile
Q11	What are the main sources or methods you use to learn about data privacy in software?	Multiple choice	RQ1
Q12	What types of personal data do you handle in your work?	Multiple choice	RQ1
Q13	Provide at least five words that come to mind when you think of privacy.	Open-ended	RQ1
Q14	I am aware of privacy laws relevant to my field, such as ISO/IEC 29100 and Privacy by Design.	Likert scale	RQ1
Q15	I have solid knowledge of privacy laws and regulations that apply to my work and how they influence software development.	Likert scale	RQ1
Q16	I have participated in internal trainings or workshops on security and privacy, including internal privacy policies and sector-specific regulations.	Likert scale	RQ1
Q17	I seek to learn about privacy on my own initiative, including reading laws and regulations and consulting specialists.	Likert scale	RQ1
Q18	I recognize the risks and concerns related to data privacy in software systems, such as data leaks, unauthorized access, and lack of user control.	Likert scale	RQ1
Q19	I know best practices and techniques to protect user privacy in software systems, including data anonymization and access control practices.	Likert scale	RQ1
Q20	I understand the difference between security and privacy, recognizing that privacy goes beyond data protection and includes aspects such as control over data and informed consent.	Likert scale	RQ1
Q21	For statements with which you strongly disagreed, describe the reason for your assessment	Open-ended	RQ1
Q22	I worry about being monitored or manipulated when using apps, social networks, or browsing the internet.	Likert scale	RQ1
Q23	I believe personal privacy is a fundamental right and we must remain alert to privacy violations.	Likert scale	RQ1

Continued on next page

ID	Question	Type	RQ
Q24	I feel personally responsible for protecting user privacy in my work as an IT professional.	Likert scale	RQ1
Q25	I perceive that many people do not care about privacy, but I believe it is important to educate and raise awareness about privacy rights.	Likert scale	RQ1
Q26	I feel frustrated with the idea that privacy is unattainable in today's digital society.	Likert scale	RQ1
Q27	I value users' informed consent before data collection and believe it is essential for building trust between users and developers.	Likert scale	RQ1
Q28	I recognize that the retention of personal data must be limited and depends on the type of data and system purpose.	Likert scale	RQ2
Q29	I consider privacy a shared responsibility across the entire IT team.	Likert scale	RQ2
Q30	For statements with which you strongly disagreed, describe the reason for your assessment	Open-ended	RQ2
Q31	Identifying privacy issues during development activities is an important part of my professional routine.	Likert scale	RQ2
Q32	When I encounter privacy issues, I escalate them to project leaders, more experienced colleagues, or security operations teams.	Likert scale	RQ2
Q33	I propose solutions to privacy issues identified during software development or operation.	Likert scale	RQ2
Q34	I have played the role of privacy advocate in my team or organization.	Likert scale	RQ2
Q35	When dealing with privacy conflicts with clients, I try to negotiate the implementation of privacy controls.	Likert scale	RQ2
Q36	I have faced suspicious client requests for excessive data collection.	Likert scale	RQ2
Q37	For statements with which you strongly disagreed, describe the reason for your assessment	Open-ended	RQ2
Q38	How often do you use privacy techniques and strategies to protect personal data in the following activities: requirements analysis, design, coding, feasibility study, application installation, deployment, testing, maintenance, operation, support.	Likert scale (matrix)	RQ2
Q39	How often do you use or have you used the following privacy strategies: encryption, minimization of personal data collection, decentralization, data sovereignty, data temporality, user control, disabling data collection, anonymization, data classification tools, code and design reviews, risk management, data flow modeling, proxy.	Likert scale (matrix)	RQ2
Q40	Do you use any other implementation strategies to ensure privacy not mentioned in the previous question?	Open-ended	RQ2
Q41	In your opinion, what is the importance level of the following privacy strategies (encryption, minimization, anonymization, etc.)?	Likert scale (matrix)	RQ2
Q42	For the strategies you considered important, describe the reason for your assessment.	Open-ended	RQ2
Q43	Regarding these strategies, how would you characterize their ease of use?	Likert scale (matrix)	RQ2
Q44	In your organization, is there a team dedicated exclusively to privacy management?	Multiple choice	RQ3
Q45	In your organization, what is the working dynamic for privacy and data protection management?	Open-ended	RQ3
Q46	In your organization, are one or more tools used for managing private data? If yes, which ones?	Multiple choice + open	RQ3
Q47	What is your perception of your organization's priority regarding privacy and data protection?	Likert scale	RQ3
Q48	In your opinion, what will be the biggest challenges for organizations in the coming years regarding practices and regulations to better protect users' privacy rights?	Open-ended	RQ3
Q49	Compliance with privacy requirements harms my productivity.	Likert scale	RQ3
Q50	I will receive incentives/recognition for adopting privacy practices.	Likert scale	RQ3
Q51	I will receive some disapproval if I do not follow privacy rules.	Likert scale	RQ3
Q52	The adoption of Privacy by Design is not compatible with my activities.	Likert scale	RQ3
Q53	My peers/leaders think I should adopt privacy practices.	Likert scale	RQ3
Q54	It is easy to become skilled in the use of privacy.	Likert scale	RQ3
Q55	Privacy improves the performance of information security in software.	Likert scale	RQ3

The questionnaire was developed and administered using the Google Forms platform. The survey remained open for two months, from July to September 2025. Participation was voluntary, and respondents were invited and encouraged to contribute through email invitations.

3.3. Pilot Study

A pilot test was conducted to evaluate the quality of the questionnaire. The form was sent to three IT staff members from a Federal Institute (IF), whose feedback was valuable for improving the survey. They suggested adjustments to the wording of some questions, the removal of redundant items, and the inclusion or modification of certain response options. Based on their feedback, the questionnaire was refined. Pilot participants took approximately 20 minutes to complete the survey, and this average time was later communicated to respondents when the survey was made publicly available. The pilot responses were not included in the data analysis.

3.4. Data Analysis

This study adopts a mixed-method approach, integrating both quantitative and qualitative research techniques. To characterize the sample and classify the most cited items by participants, percentages and graphs were employed. Additionally, correlation analysis was applied to investigate the relationships between specific variables. For the open-ended survey questions, we conducted open and axial coding following the principles of Grounded Theory [Bryant and Charmaz 2007]. Grounded Theory refers to an inductive method of generating theory from data. Studies typically include unstructured text (e.g.,

interview transcripts, field notes), but they can also encompass structured text, diagrams, images, and even quantitative data [Bryant and Charmaz 2007].

The coding process was carried out in three stages. In the first stage, two authors performed open coding of the discursive questions, segmenting the data into discrete parts and creating labels for the codes. In the second stage, the same authors conducted axial coding, reviewing the discrete parts and assigning the created codes to those segments. In the third stage, categorization and refinement of the codes took place. The coding process thus includes respondents' comments, the quotes that originated the categories, as well as the categories and subcategories created from the data.

For example, Participant R#28 provided the terms *Compliance, Consent, Security, Anonymization, and Confidentiality* when asked to list five words that come to mind when thinking about privacy (Q13). During the open coding stage, these words were segmented into discrete concepts. In the axial coding stage, *Security* was assigned to the broader category *Security and Protection*, while the other terms were distributed across complementary categories, such as *Legislation and Compliance* (Compliance, Consent), *Confidentiality, Secrecy, Anonymity* (Confidentiality, Anonymization). This example illustrates how individual responses were systematically classified into higher-level categories, as summarized in Table 3.

4. Results

This section presents the results of the survey conducted to address the research questions outlined in Section 3. We begin with an overview of the respondents' profile (Section 4.1), followed by the presentation of findings organized according to the research questions.

4.1. Respondents' Profile

Table 2 summarizes the demographic profile of the 58 respondents (Q1–Q10 from Table 1). Respondents were distributed across 15 Brazilian states plus the Federal District. The regional distribution (Q2) concentrated in the Southeast (39.6%) and Center-West (22.4%), followed by the Northeast (15.5%), North (12.1%), and South (10.4%).

The demographic data reveal a participant group composed primarily of mid-career professionals: almost half (48.3%) were between 35 and 44 years old (Q1), and the majority (72.4%) had more than 7 years of experience in IT or software development (Q7). Most respondents held advanced academic qualifications (Q3), with more than half (50.0%) having a master's degree and an additional 36.2% holding a specialization. Regarding work models (Q4), a significant portion worked on-site (44.8%), although remote (31.0%) and hybrid (24.1%) arrangements were also well represented, reflecting the diversity of work practices within IFES. In terms of roles (Q5), developers were the largest group (32.9%), followed by analysts (12.1%), security and privacy engineers (8.6%), project leaders (12.1%), and architects (6.9%), indicating coverage across the full software development lifecycle.

Considering seniority (Q6), the distribution was relatively balanced: about a quarter of the respondents (27.6%) had up to 5 years in their current role, 20.6% had between 6 and 9 years, and the remaining 51.8% had 10 or more years, evidencing a workforce with substantial accumulated experience. The vast majority were federal public servants (96.6%), aligning with the study's scope (Q8). Almost half of the respondents (43.1%)

reported having no prior experience with privacy in software, while others indicated either direct (29.3%) or indirect (27.6%) involvement (Q9), highlighting both a substantial knowledge gap and opportunities for targeted capacity building in Privacy by Design within IFES.

Age group	#	%
25–34 years	10	17.2
35–44 years	28	48.3
45–54 years	14	24.1
55–64 years	6	10.3
Region (Brazil)	#	%
Southeast (MG, SP, ES, RJ)	23	39.6
Center-West (DF, GO, MT, MS)	13	22.4
South (PR, SC, RS)	6	10.4
Northeast (BA, MA, SE, CE, AL, PI, PE, RN, PB)	9	15.5
North (AC, TO, RO, RR, AM, PA, AP)	7	12.1
Education Level	#	%
Graduated	4	6.9
Specialization	21	36.2
Master's degree	29	50.0
PhD	4	6.9
Work Model	#	%
Remote (full)	18	31.0
Hybrid (partial)	14	24.1
On-site	26	44.8
Professional Role	#	%
Developer (backend/frontend/fullstack)	19	32.9
Analyst (requirements/business)	7	12.1
Project Leader	7	12.1
Operations/Support	8	13.7
Architect (solutions, DB, etc.)	4	6.9
Security Engineer	3	5.2
Privacy Engineer	2	3.4
Others (Data Engineer, Data Scientist, Tester, IT Auditor, Database Administrator)	8	13.7
Seniority in Current Role	#	%
Up to 5 years	16	27.6
6–9 years	12	20.6
10–15 years	15	25.9
16+ years	15	25.9
Experience in IT/Software	#	%
Less than 1 year	4	6.9
1–6 years	16	27.6
7–14 years	14	24.1
15+ years	24	41.4
Employment Type	#	%
Federal Public Servant	56	96.6
Outsourced	2	3.4
Privacy Experience	#	%
No experience	25	43.1
Direct experience	17	29.3
Indirect experience	16	27.6

Table 2. Demographic profile of survey respondents (dataset, $n = 58$).

The open-ended response (Q10) revealed that most IT professionals reported having no prior experience with privacy in software, which highlights a significant knowledge and practice gap in this domain within Brazilian Federal Higher Education Institutions. Among those who indicated some level of experience, their involvement varied: a smaller group described theoretical or introductory exposure, such as awareness of institutional guidelines or participation in presentations about privacy. Others reported more concrete activities, including requirements analysis and Privacy by Design practices, or engagement in the development and maintenance of academic systems handling sensitive student

and staff data. A number of participants related privacy to their broader work in information security and IT governance, while some provided specific technical examples, such as the implementation of authentication mechanisms (e.g., OAuth, LDAP), access control, and anonymization of sensitive data. A few respondents emphasized their role in LGPD compliance projects and, in isolated cases, reported acting in formal institutional privacy roles, such as assistant data protection officer. Overall, the findings indicate that while practical and structured experience with privacy remains limited, there are scattered efforts across requirements engineering, system development, security, and compliance initiatives, suggesting potential entry points for strengthening Privacy by Design maturity in IFES.

4.2. RQ1. Perceptions and Understanding of Data Privacy

The most common sources of learning about privacy in software (Q11), as shown in Figure 1, are *documentation and guidelines* (57.9%) and *training programs or courses* (56.1%). These two categories clearly dominate, showing that practitioners tend to rely on institutional materials and structured training opportunities. Lectures (35.1%), webinars (29.8%), seminars (22.8%), and workshops (22.8%) form a second group of moderately used sources, indicating that events and collective learning opportunities also play a relevant role. More formal avenues of education, such as specialization or graduate courses, were reported by 21.1% of respondents, a frequency similar to the category "others" (21.1%). Technical tools and libraries were less cited (15.8%), while highly interactive or less conventional methods, such as hackathons (3.5%) and mentoring (3.5%), were rarely mentioned.

Overall, these findings suggest that IFES practitioners' knowledge acquisition regarding privacy is largely based on *formal documentation and structured training*, with less emphasis on experiential learning and collaborative practices. This imbalance may explain some of the gaps in practical application observed in later survey results, reinforcing the need to promote more diverse and hands-on learning strategies for Privacy by Design.

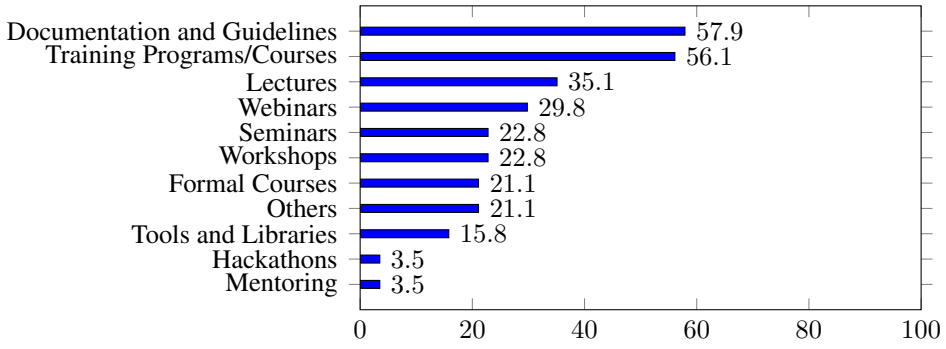

Figure 1. Sources and methods used to learn about privacy in software

Almost all respondents (96.5%) handle personally identifiable information (PII) such as names, CPF numbers, addresses, phone numbers, and emails (Q12) in their professional activities. In addition, a large majority (82.5%) reported dealing with personal characteristics (e.g., age, gender, marital status, place of birth), and nearly half (49.1%)

work with family-related data (e.g., information about relatives). More sensitive categories are also present in the respondents' context: 31.6% indicated handling sensitive data (e.g., racial or ethnic origin, religious beliefs, political opinions, genetic or biometric information) and 31.6% reported handling financial data (e.g., bank accounts, income, expenses). Less frequent but noteworthy were personal habits (12.3%), judicial or administrative records (10.5%), and psychological characteristics (8.8%), as shown in Figure 2. Overall, these results indicate that IFES professionals routinely deal with a wide spectrum of personal data—including LGPD-classified sensitive categories—reinforcing the need for systematic adoption of Privacy by Design practices to ensure compliance and mitigate exposure risks.

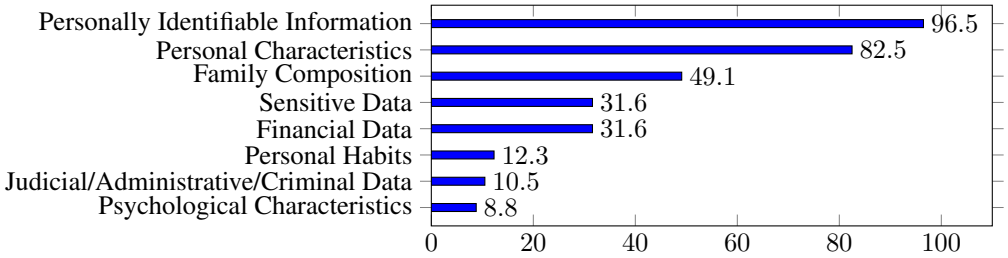

Figure 2. Types of personal data handled by respondents in their professional activities

Respondents most frequently associated privacy with *security and protection* (47 mentions), followed by *confidentiality, secrecy, and anonymity* (35). References to *personal and sensitive data* appeared in 28 responses, while *legislation and compliance* was highlighted in 25, often with explicit mention of the LGPD. Terms related to *control and access management* (22) and *ethical and social values* such as rights, freedom, and dignity (20) were also recurrent. Finally, *risks, incidents, and sanctions* (18) revealed an association between privacy and threats such as data breaches, fraud, and penalties, as shown in Table 3. These findings show that professionals perceive privacy through a predominantly technical and security-oriented lens, complemented by legal, ethical, and risk-related considerations.

Category	#	Examples of Terms
Security and Protection	47	security, protection, defense, reliability, integrity
Confidentiality, Secrecy, Anonymity	35	confidentiality, secrecy, anonymity, hidden, restriction
Personal and Sensitive Data	28	personal data, CPF, address, income, sensitive data
Legislation and Compliance	25	LGPD, law, regulation, compliance, audit
Control and Access Management	22	access, control, permission, governance, authorization
Ethical and Social Values	20	rights, freedom, citizenship, intimacy, dignity, trust
Risks, Incidents, and Sanctions	18	data breach, fraud, scam, risk, incident, fine, embarrassment

Table 3. Categories of words associated with privacy based on responses (Q13).

The results presented in Figure 3 (Q14–Q20) reveal marked differences in respondents' levels of knowledge and awareness of privacy issues. The weakest dimensions are Knowledge of Laws (Q14) and Solid Knowledge of Regulations (Q15), where the majority of responses concentrated on the negative side of the scale (over 50% between "Strongly Disagree" and "Disagree"). Similarly, Participation in Trainings (Q16) showed low engagement, with almost two-thirds of participants reporting little or no involvement in formal activities. In contrast, more favorable results were observed for Self-learning Initiative (Q17), where most respondents positioned themselves between "Agree" and

"Strongly Agree", suggesting that individual effort is the main driver of privacy learning. The strongest awareness was registered in Awareness of Risks (Q18), with over 80% of respondents agreeing or strongly agreeing that they recognize risks such as data leaks and unauthorized access. Intermediate results were found for Knowledge of Best Practices (Q19), which split between disagreement and agreement, and for Distinction Security vs Privacy (Q20), where more than 70% of respondents agreed or strongly agreed, indicating that most professionals conceptually differentiate between the two domains. Overall, these findings suggest that while IFES practitioners demonstrate strong risk awareness and a conceptual understanding of privacy distinct from security, they still lack solid regulatory knowledge and structured training opportunities.

The open-ended responses (Q21) revealed significant knowledge gaps regarding privacy laws and regulations (e.g., LGPD, ISO/IEC 29100) among the participants. Several respondents stated they had not received formal training or did not fully understand the relevant legislation. Barriers such as lack of time, motivation, and institutional support were commonly cited. Additionally, some professionals mentioned a sense of work overload, as privacy responsibilities tend to accumulate on those already in charge of information security. Despite these limitations, a few respondents indicated that they try to apply basic practices such as access control to reduce risks, even without in-depth legal or methodological knowledge. R#11 and R#18 said:

"Since I am already responsible for information security and the only one in the team trained in this area, I feel overloaded, so I tend to avoid anything related to privacy to prevent adding more responsibilities."

"I have no knowledge of the norms and legislations that define privacy rules related to my work and I have not yet found the motivation to research this topic."

A large majority of participants expressed concern about being monitored or manipulated when using digital technologies (Q22), with 57.9% strongly agreeing and 26.3% agreeing. Even more striking, 77.2% strongly agreed that personal privacy is a fundamental right (Q23), highlighting the centrality of this principle for the participants. Similarly, most respondents recognized their personal responsibility in protecting user privacy in their professional roles (Q24), with 63.2% strongly agreeing and 26.3% agreeing, and also emphasized the importance of educating and raising awareness about privacy despite perceived public apathy (Q25), with a combined 93% agreement. In contrast, perceptions about the unattainability of privacy in the digital era (Q26) were more divided, with 29.8% strongly agreeing and 33.3% agreeing, but a notable 28.1% remaining neutral, reflecting ambivalence on this point. Finally, respondents demonstrated high appreciation for informed consent as an essential foundation of trust between users and developers (Q27), with 43.9% strongly agreeing and 38.6% agreeing, as shown in Figure 3.

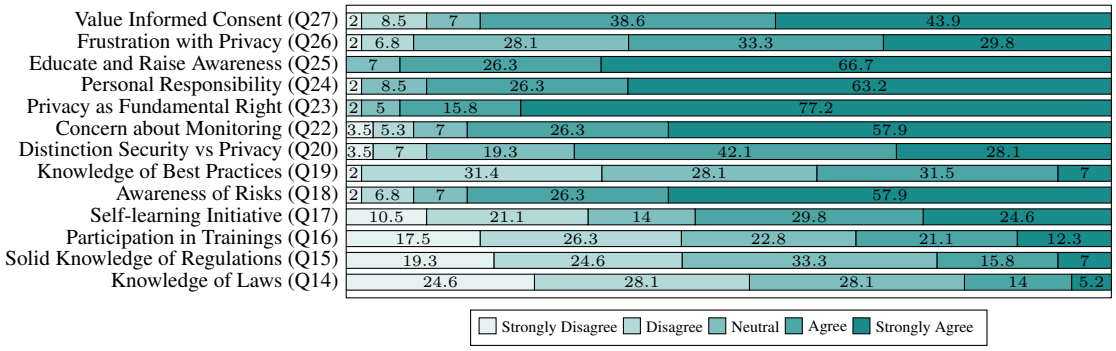

Figure 3. Knowledge, awareness, and attitudes toward privacy (Q14–Q27, $n = 58$).

RQ.1 Summary: Respondents demonstrated strong awareness of privacy as a fundamental right and a clear sense of personal responsibility, with high value placed on informed consent and risk recognition. However, significant gaps emerged in regulatory knowledge and formal training, with self-learning cited as the primary means of skill acquisition. Overall, perceptions reflected an ethical and security-oriented view of privacy, but revealed the need for broader institutional support and structured capacity building.

4.3. RQ2. Practices and Techniques for Implementing Data Privacy

There is a strong consensus among respondents regarding the principles of data retention and shared responsibility for privacy (Figure 4). For data retention (Q28), more than two-thirds (67.2%) strongly agreed and 29.3% agreed that personal data should only be retained according to its type and purpose, with minimal neutrality (3.4%) and no disagreement. Similarly, the perception of privacy as a shared responsibility (Q29) was overwhelmingly positive, with 79.3% strongly agreeing and 12.1% agreeing, while only small fractions expressed neutrality (6.6%) or disagreement (2%). These results highlight that IT professionals in IFES not only value limiting data retention but also clearly recognize privacy as a collective duty within the IT team.

Some participants reported not having completely disagreed with any statement, while others pointed out gaps in the formulation (Q30), such as the lack of questions about whether users would give up privacy in exchange for system functionalities. A recurrent point was the critique of relying exclusively on consent as a legal basis, seen as fragile and difficult to operationalize. As R#36 said:

"Consent is a fragile and difficult-to-apply instrument; instead, it is essential to adopt organizational policies, risk management, and proactive transparency measures to protect data subjects."

Responses revealed substantial variation in how practitioners address privacy in software development, as shown in Figure 4. Identifying privacy issues (Q31) was relatively common, with 44.8% of respondents agreeing or strongly agreeing, although more than one-third (36.2%) reported neutrality. A stronger pattern emerged for escalating privacy issues (Q32), where 72.5% agreed or strongly agreed, showing that reporting concerns is a well-established practice. Similarly, proposing solutions (Q33) received high agreement levels (55.1%), but a notable proportion (22.4%) remained neutral. Acting as

a privacy advocate (Q34) showed more dispersion, with responses distributed across the scale, suggesting that this role is less consistently assumed. When dealing with clients, 43.1% of participants reported negotiating privacy controls (Q35), while nearly the same proportion expressed neutrality or disagreement. Finally, facing suspicious client requests for excessive data (Q36) produced mixed results, with 43.3% agreeing or strongly agreeing, but significant portions also reporting disagreement or strong disagreement, indicating uneven exposure to such situations.

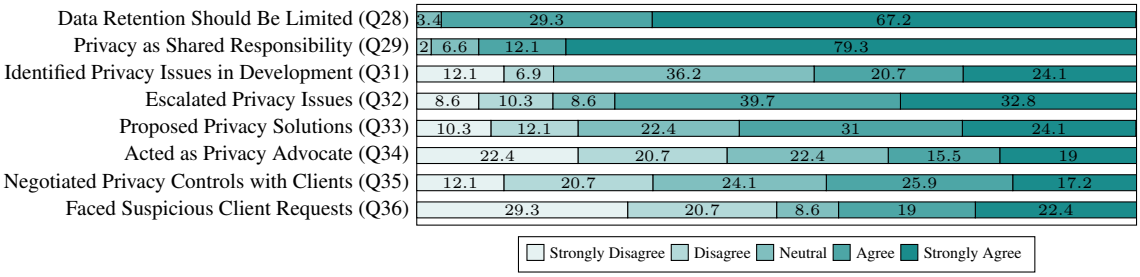

Figure 4. Perceptions about data retention, shared responsibility (Q28–Q29), and practices/experiences related to privacy in software development (Q31–Q36)

Most respondents indicated that they had not strongly disagreed with the statements because they had not experienced situations directly related to privacy practices in software projects (Q37). Several answers highlighted the limited scope of their work or lack of involvement with privacy-sensitive tasks, often delegating responsibility to designated roles such as data protection officers or information security coordinators. A few noted that the lack of experience or current responsibilities outside development influenced their answers. One respondent (R#42) pointed out that many excessive data collection requests arise not from malicious intent but from ignorance of LGPD principles:

"In general, many requests for excessive data collection occur due to lack of knowledge about the process and the LGPD. Often, when designing a form, information is included that seems useful at some future point, anticipating unnecessary collection."

The results indicate heterogeneous adoption of privacy techniques across software development activities and strategies (Figure 5). With respect to activities (Q38), respondents most frequently applied privacy measures during requirements analysis, testing, and support, while feasibility studies and coding stages showed comparatively lower use. Regarding strategies (Q39), encryption and data minimization emerged as the most consistently applied, often reported as used "frequently" or "always." In contrast, approaches such as data classification tools, code/design reviews, data sovereignty, decentralization, and proxy use registered higher rates of "never" or "rarely," highlighting a limited penetration of these practices in professional routines. Overall, the findings suggest that practitioners prioritize well-established and technically supported strategies, while less conventional or organizationally demanding approaches remain underutilized.

Few respondents reported additional strategies (Q40) beyond those listed, including monitoring and auditing tools, strict access controls, secure software development lifecycle (SDLC) practices, VPN, BitLocker, encryption, advanced authentication, and access logging.

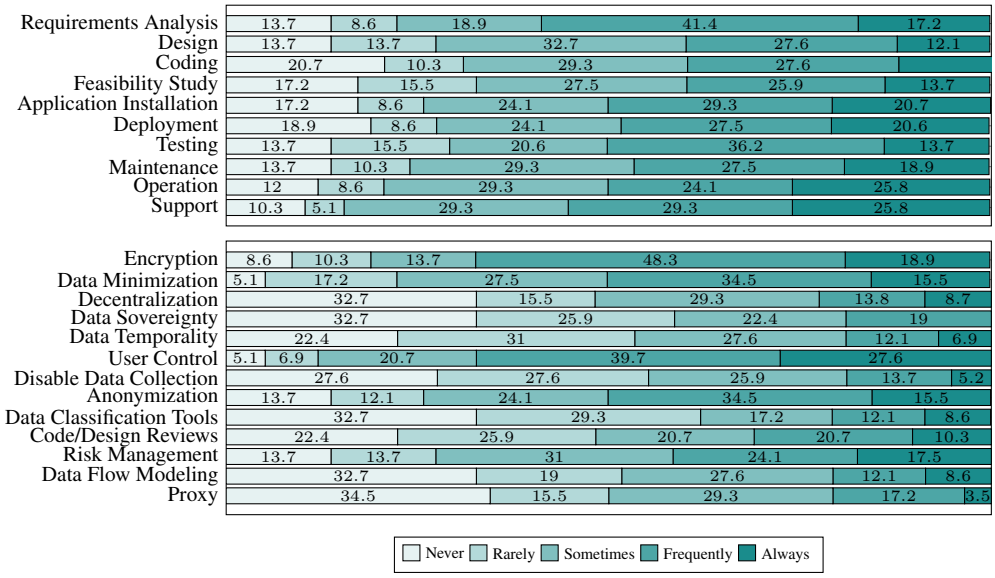

Figure 5. Frequency of using privacy techniques across development activities (top) and specific privacy strategies (bottom), $n = 58$**.**

The vast majority of respondents rated techniques such as encryption, data minimization, anonymization, risk management, and user control as "very important" (Q41), reflecting a consensus on their important role in data protection. Strategies such as decentralization, data sovereignty, and data temporality were seen as important but with greater dispersion between "important" and "moderately important" responses, suggesting more varied perceptions of their practical applicability. Although some mentions of "slightly important" or "not important" appeared in specific items (e.g., proxy), these represent a minority and reinforce that, overall, professionals strongly recognize the value of core privacy protection approaches.

Participants justified their evaluations of the most important privacy strategies (Q42) by emphasizing their central role in protecting sensitive data, ensuring compliance with legal frameworks (such as the LGPD), and reducing operational risks. Strategies such as encryption, access control, anonymization, and risk management were considered important to prevent incidents and support effective data governance. Several responses also highlighted that these practices benefit both users and organizations by fostering greater security and trust.

Regarding the perceived ease of use of privacy strategies (Q43), respondents showed a predominance of "neither easy nor difficult" evaluations across most items, indicating moderate usability overall. Core strategies such as encryption, data minimization, anonymization, and user control tended to be rated as "easy" or "very easy" by many participants, reflecting familiarity and routine application. In contrast, more complex or less common approaches—such as decentralization, data sovereignty, and proxies—attracted higher proportions of "difficult" or "very difficult," suggesting greater implementation challenges and variability in hands-on experience.

RQ.2 Summary: Respondents widely agreed on limiting data retention and recognizing privacy as a shared responsibility. Core strategies such as encryption, data minimization, anonymization, risk management, and user control were seen as very important and frequently applied. In contrast, approaches like decentralization, data sovereignty, temporality, and proxies were less common and considered more difficult to implement, indicating variability in hands-on experience and practical applicability.

4.4. RQ3.Factors Influencing Privacy Adoption and Maturity

Responses to Q44 revealed a fragmented scenario across Brazilian Federal Higher Education Institutions. Only 17% of respondents reported the existence of a dedicated internal privacy team. In contrast, the majority (83%) indicated the absence of such teams, with answers distributed among different alternatives: responsibility being distributed among IT teams (21%), concentrated in the DPO and their supporting staff (23%), or acknowledging that no formal role exists (19%). Additionally, a significant portion (21%) answered that they did not know whether such a team existed in their institution, reflecting a lack of institutional communication and visibility regarding privacy governance.

The findings indicate that dedicated privacy governance structures remain an exception rather than the rule within IFES. In most cases, responsibility is either dispersed across IT staff or concentrated in the DPO, which may limit institutional capacity to systematically integrate Privacy by Design principles. The high rate of respondents reporting a lack of knowledge about the existence of such teams further suggests weak communication and transparency regarding privacy responsibilities in these organizations.

In relation to the organizational dynamics for privacy and data protection (Q45), participants' responses reveal a fragmented scenario. A substantial number of respondents reported having no knowledge of existing practices or stated that no formal structure is in place. Others indicated that responsibility is distributed across IT teams or delegated to the Data Protection Officer (DPO), often without adequate staff or resources. Several mentioned the existence of committees or commissions, but described them as largely formalistic or with limited effectiveness. In many cases, privacy is handled reactively, typically in response to incidents or external demands. On the other hand, some respondents highlighted the presence of institutional guidelines and policies (e.g., LGPD, POSIC), technical controls such as access restrictions, data minimization, and monitoring, as well as efforts to raise user awareness. A smaller group described initiatives to embed privacy into software development processes, particularly during requirements analysis and access permission design, as shown in Table 4. Overall, the findings indicate that while there are emerging practices, most institutions still lack systematic and proactive approaches to privacy management, relying instead on fragmented responsibilities, reactive measures, and limited organizational capacity.

The majority of participants reported having no knowledge of specific tools or stated explicitly that none are used in their organizations for managing personal data (Q46), as shown in Table 5. This highlights a widespread lack of visibility or absence of structured technological support for privacy management across IFES. Only a few respondents mentioned concrete tools or initiatives: OneTrust, TrustArc, and DPOnet (in

Table 4. Categories of organizational dynamics for privacy and data protection (Q45, open-ended responses).

Category	#	Examples (translated)
No knowledge / Cannot describe	14	"I do not know the dynamics of privacy and data protection in my organization."
No formal structure	11	"There is no defined dynamic yet." / "There is no formal role or team."
Shared responsibility (distributed across IT teams)	9	"Responsibilities are shared among IT teams." / "Each team deals with it separately, without a protocol."
DPO responsibility (with or without support)	8	"The DPO and their staff are in charge." / "There is a DPO named, but no support team."
Committees or commissions	6	"There is a committee that is activated when a vulnerability is found." / "A commission was appointed but with little productivity."
Reactive / On-demand approach	7	"It is reactive, handled only when incidents occur." / "TI responds to demands when requested."
Internal policies and guidelines (POSIC, LGPD)	8	"Follow institutional POSIC." / "Based on LGPD and institutional policies."
Technical controls (access, minimization, backups, monitoring)	7	"Adoption of minimum necessary principle in each system." / "Authentication, access control, monitoring of systems."
Integration with software processes (requirements, permissions, dev cycle)	4	"We evaluate privacy points in requirements gathering and module access permissions."
Awareness and training efforts	3	"Annual training courses are offered." / "User awareness activities are being developed."

pilot tests), as well as IBM Guardium and Varonis for Data Loss Prevention (DLP) and monitoring. Others referred to more general mechanisms, such as institutional policies, authentication via LDAP, or data protection guidelines, rather than specialized privacy management platforms. A single mention of SECURAMDATA also appeared. Overall, the responses indicate that while a handful of institutions are experimenting with recognized privacy management solutions, most professionals either are unaware of such tools or work in environments where no dedicated platforms are formally adopted. This gap reinforces the reliance on policies and manual practices, suggesting low maturity in the technological dimension of Privacy by Design adoption.

Table 5. Categories of responses regarding the use of tools for managing personal data (Q46).

Category	#	Examples of Terms/Responses
No knowledge/None	46	"I have no knowledge", "None", "There is no"
Specialized Tools (Privacy/DLP)	6	OneTrust, TrustArc, DPOnet (pilot), IBM Guardium, Varonis (DLP/monitoring), SECURAMDATA
Institutional Policies / General Mechanisms	2	Authentication (LDAP), institutional data protection policies
Restricted/Unclear Information	1	"Restricted information"

Regarding organizational priority on privacy and data protection (Q47), most participants perceived it as moderate (36.2%), followed by perceptions of high priority (22.4%) and low priority (17.2%). A considerable proportion of respondents also viewed privacy as a very low priority (13.8%), while only a small minority (10.3%) considered it a very high priority. These results indicate that, although awareness of privacy is growing, the majority of IT professionals in Federal Higher Education Institutions still perceive it as not being treated as a strategic priority in their organizations (Figure 6).

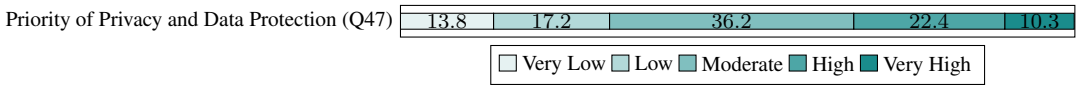

Figure 6. Perceived organizational priority regarding privacy and data protection (Q47, $n = 55$).

In relation to the perceived challenges for organizations in the coming years regarding privacy practices and regulations (Q48), participants highlighted a wide range of concerns. The most frequently mentioned category refers to **technological challenges and the growing use of Artificial Intelligence**, cited in 16 responses. These include the risks posed by generative AI and machine learning systems, often seen as "black boxes", the difficulty of ensuring transparency and explainability, as well as the use of massive

training datasets that may contain personal or sensitive information. Respondents also emphasized the intensification of cyberattacks and the exponential growth of data volumes as critical threats.

The second most recurrent theme is **organizational culture and awareness**, with 14 mentions. Participants reported the absence of consolidated strategies, a predominantly reactive approach to privacy, and the lack of prioritization of the topic by senior management. Several answers stressed that without a cultural shift and recognition of privacy as a strategic concern, technical measures alone will not suffice. **Human and financial resources** were also identified as a major barrier (11 mentions). Respondents pointed to the shortage of skilled professionals, budgetary restrictions, and the overload of existing IT staff, who often accumulate multiple responsibilities without the necessary support to implement privacy-oriented measures.

Another relevant challenge is the **regulatory complexity and compliance landscape** (10 mentions). The constant evolution of national and international privacy frameworks such as the LGPD and GDPR, the ambiguity of legal requirements, and tensions between transparency in public institutions and individual privacy rights were described as difficult to reconcile in practice. Finally, **legacy systems and technical debt** (7 mentions) were highlighted as obstacles, given that many existing applications were not designed with Privacy by Design principles. Retrofitting these systems to comply with current regulations is seen as costly and complex, demanding significant effort to adapt infrastructure and processes. Overall, the findings (Table 6) show that organizations face intertwined technological, cultural, human, regulatory, and infrastructural challenges. Addressing these demands will require not only compliance with evolving laws but also investments in people, governance, and proactive strategies that integrate privacy into technological innovation.

Category	#	Examples of Terms / Quotes
Technological Challenges and AI	16	generative AI, machine learning, black-box models, training data, cyberattacks, increasing data volumes, transparency issues
Organizational Culture and Awareness	14	lack of prioritization, reactive approach, lack of awareness, absence of strategy, culture of compliance, "management must recognize importance"
Human and Financial Resources	11	lack of skilled workforce, resource scarcity, need for investment, staff overload, insufficient budget
Regulatory Complexity and Compliance	10	evolving regulations, LGPD, GDPR, unclear guidelines, conflicts between transparency and privacy, legal enforcement, judiciary interventions
Legacy Systems and Technical Debt	7	old systems without PbD, retrofitting costs, technical debt, difficulty adapting existing infrastructure

Table 6. Categories of future challenges for privacy practices and regulations based on participants' responses (Q48).

In relation to perceptions of privacy practices (Q49–Q55), the results reveal a heterogeneous picture (Figure 7). Regarding productivity (Q49), almost one-third of respondents expressed neutrality (34.5%), while 22.4% disagreed and 20.7% strongly disagreed with the idea that privacy requirements harm their productivity. Nevertheless, a considerable portion still perceived negative impacts, with 12.1% agreeing and 10.3% strongly agreeing, suggesting that productivity concerns persist in some contexts. When asked about incentives and recognition for adopting privacy practices (Q50), the majority strongly disagreed (43.1%) or disagreed (22.4%), and only 15.5% indicated agreement (8.6%) or strong agreement (6.9%). This shows that most professionals perceive a lack of institutional rewards or recognition for engaging in privacy-related activities.

A different pattern emerged in terms of accountability (Q51). Almost half of the respondents agreed (15.5%) or strongly agreed (19.0%) that they would face disapproval

if they did not follow privacy rules, while 34.5% remained neutral. These results suggest that, although incentives are scarce, organizational expectations and peer accountability mechanisms may still enforce compliance. Perceptions about the compatibility of Privacy by Design with work activities (Q52) were predominantly neutral (43.1%), with a significant number of participants disagreeing (19.0%) or strongly disagreeing (20.7%). Only a small minority (17.2%) agreed or strongly agreed that PbD is not compatible with their tasks, indicating that most professionals do not reject the applicability of PbD, but may remain uncertain about its integration (Figure 7).

Social influence (Q53) appears to play a stronger role: 24.1% agreed and 20.7% strongly agreed that peers or leaders expect them to adopt privacy practices, although neutrality was still high (34.5%). This highlights the importance of organizational culture and leadership in shaping professional behavior. With respect to skill acquisition (Q54), almost 39.7% of participants remained neutral, and 31.0% disagreed that it is easy to become skilled in privacy. Only 12.0% agreed or strongly agreed, suggesting a widespread perception of difficulty or lack of opportunities to develop expertise in privacy practices. Finally, participants strongly associated privacy with improved information security performance (Q55). Half of the respondents strongly agreed (50.0%) and another 29.3% agreed, while only marginal shares expressed neutrality (15.5%) or disagreement (5.2%). This indicates broad consensus that privacy is not only complementary to security but also a driver of stronger protection in software development (Figure 7).

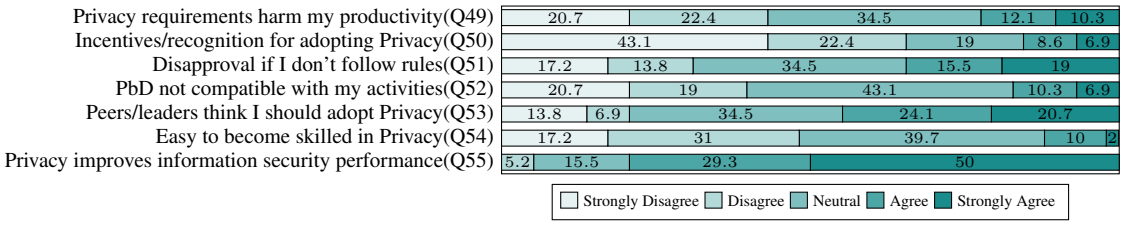

Figure 7. Perceptions of privacy practices and impacts (Q49–Q55)

RQ.3 Summary: Findings indicate fragmented organizational structures for privacy, with most institutions lacking dedicated teams or tools and often adopting reactive approaches. While professionals recognize privacy's importance, barriers such as limited training, unclear responsibilities, and scarce resources hinder maturity. Nonetheless, strong social influence and consensus on privacy's positive impact on information security suggest favorable conditions for advancing Privacy by Design adoption in IFES.

5. Discussion

The findings of this study confirm that, although IT professionals in Brazilian Federal Higher Education Institutions (IFES) demonstrate awareness of privacy as a fundamental right and recognize its shared responsibility, the adoption of systematic Privacy by Design (PbD) practices remains limited. The results highlight gaps in regulatory knowledge, lack of structured training, and uneven adoption of privacy-enhancing strategies. These observations resonate with and extend the insights provided by prior studies.

First, the predominance of informal learning and self-study as the main avenues for acquiring privacy knowledge reinforces the immaturity of the field described by Spósito et al. [Spósito et al. 2025]. Their review revealed that most proposed methods and tools for privacy requirements engineering remain confined to academic contexts, with limited industrial adoption. Our survey provides empirical confirmation of this limitation, showing that even within IFES, structured training is rare and practical knowledge about privacy frameworks is insufficient.

Second, the persistent challenges in translating legal and regulatory requirements into daily practices directly echo the 18 categories of compliance difficulties identified by Rocha and Canedo [Rocha et al. 2023]. The lack of clear methodologies, standardized tools, and institutional support reported by our respondents mirrors these broader difficulties, suggesting that IFES face the same systemic obstacles observed across organizations worldwide. Our results contribute concrete empirical evidence of how such challenges manifest in public-sector software development.

Third, the findings corroborate the socio-technical perspective emphasized by Matos et al. [Matos et al. 2025], who highlighted how developers often conflate security and privacy and how organizational culture influences the integration of PbD. In our survey, respondents revealed both conceptual confusions and institutional cultures that relegate privacy to secondary importance, reinforcing the argument that advancing PbD maturity requires addressing organizational and behavioral dynamics, not only technical guidelines.

Fourth, the uneven adoption of specific strategies—such as anonymization, encryption, and minimization being prioritized while decentralization or sovereignty are rarely used—points to the need for practical instruments to guide decision-making. This aligns with Andrade et al. [Andrade et al. 2024], who mapped PbD principles to Privacy Patterns as a way to bridge abstract concepts with concrete engineering practices. Our results indicate that such instruments could be particularly useful in IFES, where practitioners expressed uncertainty about advanced strategies and relied heavily on basic, familiar techniques.

Finally, the overall immaturity of privacy practices in IFES resonates strongly with the multiple case studies reported by Andrade et al. [Andrade et al. 2023], who observed that organizations typically address privacy only in later stages of development, lack dedicated specialists, and rarely use specialized tools. Our survey extends these observations by providing quantitative evidence from a national sample of IT professionals in higher education, demonstrating that privacy is still treated reactively and often subordinated to other institutional priorities.

In summary, the results of this study complement and extend previous research by empirically showing how the gaps, barriers, and cultural factors identified in the literature persist in the specific context of IFES. The findings strengthen the argument that advancing PbD maturity requires not only technical frameworks but also instruments and organizational measures that promote structured training, clarify responsibilities, and foster a proactive privacy culture in public-sector software development.

6. Threats to Validity

Following Wohlin et al. [Wohlin et al. 2012], the threats to the validity of this study are discussed in four categories: construct, internal, conclusion, and external validity. Mitigation strategies adopted during the research process are also presented. **Construct validity** refers to the extent to which the survey instrument accurately captures the concepts of interest, in this case, perceptions and practices related to Privacy by Design (PbD). To mitigate this threat, the questionnaire was carefully designed based on established frameworks (e.g., LGPD, ISO/IEC 29100, PbD principles) and validated in a pilot study with three IT professionals from Federal Institutes. Their feedback allowed us to refine the wording, remove redundant items, and ensure that the questions were understandable and aligned with the research objectives. Additionally, a mix of closed-ended Likert scales and open-ended questions was used to capture both quantitative and qualitative insights, enhancing construct coverage.

Internal Validity concerns potential factors that may bias the responses or introduce confounding variables. Since this study was based on a survey, typical threats include misunderstanding of the questions, respondent fatigue, or socially desirable answers. We mitigated these issues by ensuring anonymity of participation, which encouraged honest responses, and by limiting the estimated completion time to approximately 20 minutes, reducing fatigue. The invitation emphasized the voluntary nature of participation, minimizing pressure or coercion. Screening questions were also used to ensure only IT professionals from Brazilian Federal Higher Education Institutions (IFES) participated. **Conclusion validity** relates to the robustness of the analysis and the reliability of conclusions. To mitigate this, we applied both descriptive and inferential statistical techniques (e.g., correlation analysis) complemented by qualitative coding using Grounded Theory for open-ended questions. Triangulating quantitative and qualitative findings enhanced the reliability of our interpretations. Moreover, percentages and graphical visualizations were provided to increase transparency and replicability of the results. **External validity** concerns the generalizability of the findings beyond the studied sample. The survey reached 58 IT professionals across 15 Brazilian states, ensuring geographic and institutional diversity. However, as participation was voluntary, self-selection bias cannot be entirely ruled out, and the sample may not represent all IFES professionals. To mitigate this limitation, the results were interpreted as indicative rather than definitive, and compared with findings from related studies in the literature (e.g., [Matos et al. 2025, Rocha et al. 2023]). This comparative approach provides additional support for the general relevance of the identified gaps and practices.

7. Conclusion

This study investigated the perceptions, practices, and organizational factors that influence the adoption of Privacy by Design (PbD) in the context of Brazilian Federal Higher Education Institutions (IFES). Through a mixed-method survey with 58 IT professionals across 15 states, we provided empirical evidence on how privacy is understood and implemented in software development processes within the public higher education sector.

The results indicate a strong recognition of privacy as a fundamental right and a shared responsibility among IT teams. Respondents demonstrated awareness of core principles such as data retention limits, informed consent, and the use of strategies like

encryption, minimization, and anonymization. However, significant gaps remain in regulatory knowledge, structured training, and the systematic adoption of advanced privacy-enhancing practices. Organizational structures are fragmented, with most institutions lacking dedicated privacy teams or tools, and privacy often being addressed reactively rather than proactively. While professionals acknowledged that privacy strengthens information security, they also perceived barriers such as limited incentives, scarce resources, cultural resistance, and the complexity of retrofitting legacy systems.

These findings complement and extend related work in privacy engineering and requirements engineering. As highlighted by Spósito et al. [Spósito et al. 2025], research on privacy requirements remains largely academic, with limited industrial adoption, a limitation that our survey empirically confirmed within IFES. Rocha and Canedo [Rocha et al. 2023] identified compliance challenges such as legal ambiguity and lack of standardized tools; our results illustrate how these obstacles manifest concretely in public-sector organizations. In line with Matos et al. [Matos et al. 2025], we observed that misconceptions between privacy and security, combined with weak organizational cultures, hinder effective PbD adoption. Furthermore, our findings reinforce the argument of Andrade et al. [Andrade et al. 2024, Andrade et al. 2023] on the need for practical instruments and empirical validation, as most IFES professionals reported difficulty integrating advanced strategies and treating privacy systematically throughout the software lifecycle.

From a practical perspective, the study suggests that advancing privacy maturity in IFES requires a combination of actions: (i) structured and continuous training programs for IT professionals, (ii) development of institutional frameworks that clarify roles and responsibilities, (iii) adoption of dedicated tools to support privacy management, and (iv) cultural transformation that embeds privacy as a strategic value in academic and administrative systems. From a research perspective, the study contributes a diagnostic instrument that exposes maturity gaps and socio-technical barriers, offering a foundation for future frameworks and empirical evaluations of PbD adoption in the public sector.

In conclusion, while progress has been made in raising awareness and applying core practices, privacy in IFES remains at an early stage of maturity. Addressing the identified gaps requires not only compliance with the LGPD but also systematic efforts to embed Privacy by Design into the culture, processes, and technologies of public higher education institutions. By advancing this agenda, IFES can strengthen both legal compliance and the trust of the academic community, while contributing to the broader development of privacy engineering in practice.

Data Availability Statement

The data that support the findings of this study are openly available in Zenodo at `https://zenodo.org/records/17247244`.

Acknowledgments

This study was financed in part by the Project No. TED 33/2023 “Pesquisa Aplicada em Privacidade e Segurança da Informação na Diretoria de Privacidade e Segurança da Informação da Secretaria de Governo Digital” – Diretoria de Privacidade e Segurança da Informação (DPSI)/Centro de Excelência em Privacidade e Segurança (CEPS)/ Secretaria de Governo Digital (SGD) do Ministério da Gestão e da Inovação (MGI) em Serviços

Públicos do Governo Federal; and Conselho Nacional de Desenvolvimento Científico e Tecnológico CNPq (Grant 300883/2025-0).

References

Al-Slais, Y. (2020). Privacy engineering methodologies: A survey. In *2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT)*, pages 1–6.

Alexey Andreev (2025). Rockyou2024 e as outras quatro maiores violações de dados da história. https://www.kaspersky.com.br/blog/top-five-data-breaches-in-history/22917/.

Alves, C. and Neves, M. (2021). Especificação de requisitos de privacidade em conformidade com a LGPD: resultados de um estudo de caso. In de Menezes Cruz, M. L. P., Hadad, G. D. S., and Marques, J. C., editors, *Anais do WER21 - Workshop em Engenharia de Requisitos, Brasilia, BSB, Brasil, August 23-27, 2021*. Editora PUC-Rio.

Andrade, V. C., Reinehr, S. S., Freitas, C. O. A., and Malucelli, A. (2023). Personal data privacy in software development processes: A practitioner's point of view. In Hu, J., Min, G., Wang, G., and Georgalas, N., editors, *22nd IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2024, Exeter, UK, November 1-3, 2023*, pages 2727–2734. IEEE.

Andrade, V. C., Ribeiro, R. D., dos Passos Canteri, R., Reinehr, S. S., de A. Freitas, C. O., and Malucelli, A. (2024). Privacy in practice: Exploring concrete relationships between privacy patterns and privacy by design principles in software engineering. In OliveiraJr, E., de Guzmán, I. G. R., Ayala, C. P., Murta, L., Barcelos, M. P., Brito, I. S. S., Neto, A., Valderas, P., Paludo, M., Reinehr, S. S., Malucelli, A., and Cruz-Lemus, J. A., editors, *27th Iberoamerican Conference on Software Engineering, CIbSE 2024, Curitiba, Brazil, May 6-10, 2024*, pages 271–285. Curran Associates.

André Luiz Dias Gonçalves (2025). Tudo sobre o vazamento de dados de 223 milhões de brasileiros. https://www.tecmundo.com.br/software/210168-tudo-vazamento-dados-223-milhoes-de-brasileiros.htm.

Bryant, A. and Charmaz, K. (2007). *The Sage handbook of grounded theory*. Sage, https://uk.sagepub.com/en-gb/eur/the-sage-handbook-of-grounded-theory/book234413.

Bu, F., Wang, N., Jiang, B., and Jiang, Q. (2021). Motivating information system engineers' acceptance of privacy by design in china: An extended UTAUT model. *Int. J. Inf. Manag.*, 60:102358.

Bu, F., Wang, N., Jiang, B., and Liang, H. (2020). "privacy by design" implementation: Information system engineers' perspective. *Int. J. Inf. Manag.*, 53:102124.

Canedo, E. D., Calazans, A. T. S., Masson, E. T. S., Costa, P. H. T., and Lima, F. (2020). Perceptions of ICT practitioners regarding software privacy. *Entropy*, 22(4):429.

Cavoukian, A. (2012). Privacy by design [leading edge]. *IEEE Technol. Soc. Mag.*, 31(4):18–19.

Chander, A. and Land, M. (2014). United nations general assembly resolution on the right to privacy in the digital age. *International Legal Materials*, 53(4):727–731.

Confessore, N. (2018). Cambridge analytica and facebook: The scandal and the fallout so far. *The New York Times*, 4:2018.

de Chaves, S. A. and Benitti, F. B. V. (2023). Privacy by design in software engineering: An update of a systematic mapping study. In Hong, J., Lanperne, M., Park, J. W., Cerný, T., and Shahriar, H., editors, *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing, SAC 2023, Tallinn, Estonia, March 27-31, 2023*, pages 1362–1369. ACM.

Ferrão, S. É. R., Silva, G. R. S., Canedo, E. D., and Mendes, F. F. (2024). Towards a taxonomy of privacy requirements based on the LGPD and ISO/IEC 29100. *Inf. Softw. Technol.*, 168:107396.

Hadar, I., Hasson, T., Ayalon, O., Toch, E., Birnhack, M., Sherman, S., and Balissa, A. (2018). Privacy by designers: software developers' privacy mindset. *Empir. Softw. Eng.*, 23(1):259–289.

Matos, A., Patrício, M., Nicolau, M. I., Canedo, E. D., Pereira, J. A., and Uchôa, A. G. (2025). Data privacy in software practice: Brazilian developers' perspectives. *J. Internet Serv. Appl.*, 16(1):299–319.

Menegazzi, D. and Silva, C. (2023). Conformidade com a LGPD por meio de requisitos de negócio e requisitos de solução. In Antonelli, L., Lucena, M., and Portugal, R. L. Q., editors, *Anais do WER23 - Workshop em Engenharia de Requisitos, Porto Alegre, RS, Brasil, August 15-17, 2023*. LFS (UFRN, Brasil).

Morales-Trujillo, M., Matla-Cruz, E., García-Mireles, G., and Piattini, M. (2018). A systematic mapping study of privacy by design. *Avances en Ingenieria de Software a Nivel Iberoamericano, CIbSE*, 22(1):107–120.

Presidência da República do Brasil (2018). Lei nº 13.709, de 14 de agosto de 2018. `https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm`.

Ribak, R. (2019). Translating privacy: Developer cultures in the global world of practice. *Information, Communication & Society*, 22(6):838–853.

Rocha, L. D., Silva, G. R. S., and Canedo, E. D. (2023). Privacy compliance in software development: A guide to implementing the LGPD principles. In Hong, J., Lanperne, M., Park, J. W., Cerný, T., and Shahriar, H., editors, *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing, SAC 2023, Tallinn, Estonia, March 27-31, 2023*, pages 1352–1361. ACM.

Rodrigues, G. A. P., Fernandes, P. A. G., Serrano, A. L. M., Filho, G. P. R., Vergara, G. F., Bispo, G. D., de Oliveira Albuquerque, R., and Gonçalves, V. P. (2025). From rockyou to rockyou2024: Analyzing password patterns across generations, their use in industrial systems and vulnerability to password guessing attacks. *J. Internet Serv. Appl.*, 16(1):69–86.

Rodrigues, G. A. P., Serrano, A. L. M., Lemos, A. N. L. E., Canedo, E. D., de Mendonça, F. L. L., de Oliveira Albuquerque, R., Orozco, A. L. S., and García-Villalba, L. J.

(2024). Understanding data breach from a global perspective: Incident visualization and data protection law review. *Data*, 9(2):27.

Sangaroonsilp, P., Dam, H. K., Choetkiertikul, M., Ragkhitwetsagul, C., and Ghose, A. (2023). A taxonomy for mining and classifying privacy requirements in issue reports. *Inf. Softw. Technol.*, 157:107162.

Secretaria de Governo Digital (28 de março de 2023). Portaria nº 852 sgd/mgi. https://www.gov.br/funai/pt-br/centrais-de-conteudo/publicacoes/relatorios/legislacao-de-ouvidoria/l-f-portaria_sgd_mgi_852_28-3-2023_ppsi.pdf.

Spiekermann, S. (2012). The challenges of privacy by design. *Commun. ACM*, 55(7):38–40.

Spósito, S. L., Targino, J. F. G., Silva, G. R. S., Peotta, L., Porto, D. d. P., Mendonça, F. L. L., and Canedo, E. D. (2025). A comprehensive review of techniques, methods, processes, frameworks, and tools for privacy requirements. *Journal of Internet Services and Applications*, 16(1):508–529.

Tribunal de Contas da União (2025). Tcu verifica risco alto à privacidade de dados pessoais coletados pelo governo. https://portal.tcu.gov.br/imprensa/noticias/tcu-verifica-risco-alto-a-privacidade-de-dados-pessoais-coletados-pel

United Nations (2025). Universal declaration of human rights at 70: 30 articles on 30 articles - article 12. https://www.ohchr.org/en/press-releases/2018/11/universal-declaration-human-rights-70-30-articles-30-articles-article

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2012). *Experimentation in Software Engineering*. Springer.