Journal of Software Engineering Research and Development, 2026, 14:11, doi: 10.5753/jserd.2026.7125

Privacy by Design Maturity in Software Development: An Empirical Study in Brazilian Federal Higher Education Institutions

Fernando Elias de Oliveira [**University of Brasília (UnB)** | ***fernandoeliasti@gmail.com***]
Stefano Luppi Spósito [**University of Brasília (UnB)** | ***stefanoluppi@hotmail.com***]
Fabiana Freitas Mendes [**Massey University** | ***f.mendes@massey.ac.nz***]
Edna Dias Canedo [**Federal University of Pernambuco (UFPE), Computer Science Center (CIn)** | ***ednacanedo@cin.ufpe.br***]

Abstract

Context: Privacy by Design (PbD) has emerged as a key approach for embedding data protection into software systems from their inception, a pressing concern in Brazil after the enactment of the General Data Protection Law (LGPD). However, public organizations such as Brazilian Federal Higher Education Institutions (IFES) still struggle to translate legal requirements into concrete engineering practices and governance routines. **Goal:** This study investigates how IT professionals in IFES perceive, adopt, and operationalize PbD in software development, and which organizational, technical, and individual factors influence the maturity of privacy practices in this context. **Method:** We conducted a mixed-method survey with 58 IT professionals from IFES across 15 Brazilian states. The instrument combined 46 closed-ended and 9 open-ended questions covering privacy knowledge, attitudes, behaviors, strategies, and organizational conditions. **Results:** Respondents strongly recognize privacy as a fundamental right and frequently handle personal and sensitive data in their daily work. Core strategies such as encryption, data minimization, anonymization, risk management, and user control are perceived as very important and are more often applied, whereas decentralization, data sovereignty, and temporality remain uncommon and are more difficult to implement. Inferential analysis using Spearman's rank correlation indicates that the adoption of privacy strategies is significantly associated with perceived importance ($\rho = 0.37$, $p = 0.005$), while organizational, social, and usability-related factors show no statistically significant associations. **Conclusion:** PbD maturity in IFES remains incipient, characterized by fragmented and reactive practices driven more by individual perceptions than by institutional structures. Advancing this maturity requires structured training, clearer roles and responsibilities, better tool support, and strategies that reinforce the perceived importance of privacy among practitioners.

Keywords: *Privacy by Design, Privacy Engineering, Software Development Lifecycle, LGPD, Public Sector*

1 Introduction

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

The rapid evolution of the internet and digital ecosystems has intensified the scale of personal data collection and profiling (Spósito et al., 2025b). This scenario underscores the need for rigorous privacy preserving techniques in information technology processes, enabling society to benefit from digital services securely (Spósito et al., 2025b; Chander and Land, 2014). High profile incidents such as the Cambridge Analytica scandal illustrate how personal data can be exploited for commercial, political, or strategic purposes (Confessore, 2018).

Recent large scale breaches demonstrate the growing severity of privacy risks. In 2024, the RockYou2024 leak exposed nearly 10 billion user credentials (Rodrigues et al., 2025; Alexey Andreev, 2025). Major platforms have repeatedly suffered attacks: 533 million Facebook users were affected in 2019, and Yahoo experienced the exposure of over 3 billion accounts (Alexey Andreev, 2025). In Brazil, more than 223 million citizens had sensitive data leaked in 2021 credit scores, addresses, and taxpayer IDs, with strong indications that the source was SERASA (Gonçalves, André Luiz Dias, 2025). Data breaches in the Brazilian government alone exceeded 9,000 incidents in 2024, more than 20 times the number recorded in 2020 (Rodrigues et al., 2024). Although these platforms collect data to provide services, repeated exposures have left users highly vulnerable (Spiekermann, 2012).

Safeguarding privacy in organizational information systems is a complex socio-technical challenge involving multiple domains and stakeholders (Matos et al., 2025). New technologies must incorporate mechanisms that prevent or mitigate risks to the fundamental right of data protection, as guaranteed by the Brazilian Constitution (Alves and Neves, 2021). Cavoukian (Cavoukian, 2012), a pioneer of the concept of Privacy by Design (PbD), argues that privacy must be embedded directly into system design rather than treated as an afterthought. PbD is grounded in seven principles proactivity, default protection, embedded design, full functional-

ity, end-to-end security, transparency, and respect for user privacy (de Chaves and Benitti, 2023) and offers a structured way to integrate privacy into engineering artifacts and processes Rocha and Canedo (2025).

In Brazil, the General Data Protection Law (LGPD) establishes strict rules for the processing of personal data (Presidência da República do Brasil, 2018). Within this regulatory context, PbD provides a methodological foundation for embedding data minimization, purpose limitation, and security-by-design into software systems. However, audits by the Federal Court of Accounts (TCU) reveal that Federal Higher Education Institutions (IFES) and oversight bodies lack consolidated methods to assess or monitor privacy maturity (Tribunal de Contas da União, 2025). This motivates the following research question: **RQ. What is the level of maturity in the implementation of Privacy by Design principles within information systems of Brazilian Federal Higher Education Institutions?** Addressing this question requires empirical evidence on how privacy practices are operationalized in real software processes.

The increasing digitalization of services provided by IFES, reinforced by Ordinance SGD/MGI No. 852 of March 28, 2023 (Secretaria de Governo Digital, 2023), amplifies the need for robust privacy practices aligned with the LGPD Presidência da República do Brasil (2018). Nonetheless, studies indicate persistent difficulties in LGPD compliance, often attributed to limited methodologies, fragmented governance, and technical and organizational constraints (Rocha et al., 2023). This study contributes to filling this gap by assessing how PbD practices are adopted in IFES and by providing evidence to support the evaluation and improvement of privacy maturity in software processes.

The lack of systematic approaches to integrate privacy into development activities may compromise the protection of highly sensitive institutional data, including academic records, financial data, medical information, and research participation records. Without applying principles such as minimization, access control, purpose limitation, and security-by-design early in the software lifecycle, these data may be exposed or misused (Kuliamin et al., 2025; Menegazzi and Silva, 2023). Such gaps not only pose legal risks under the LGPD but also threaten the integrity of core institutional missions, including teaching, research, and administration. Embedding privacy engineering into software processes is thus a strategic requirement for regulatory compliance, risk mitigation, and institutional resilience (Bu et al., 2020).

TCU audits indicate that over 70% of IFES exhibit deficiencies in information security and privacy practices, exposing sensitive academic and socioeconomic data of students and staff (Tribunal de Contas da União, 2025). Prior incidents have already caused reputational harm and increased the likelihood of sanctions under the LGPD, which authorizes 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 becomes necessary to mitigate legal risks, preserve community trust, and ensure continuity of institutional activities.

In light of these challenges, this study provides an empirical assessment of how Privacy by Design is understood and operationalized within IFES. The results reveal that, although IT professionals strongly recognize privacy as a fundamental right and routinely handle sensitive data, PbD maturity remains incipient. Regulatory knowledge and formal training are limited, dedicated privacy teams and tools are rare, and privacy is often treated as a moderate or low institutional priority. While core strategies such as encryption, minimization, anonymization, and risk management are widely valued and frequently applied, more advanced approaches such as decentralization, data sovereignty, and temporality remain uncommon and difficult to implement. In addition to descriptive findings, this study also investigates whether individual, organizational, and contextual factors are statistically associated with the adoption of Privacy by Design practices. This enables an exploratory understanding of potential drivers of privacy adoption beyond descriptive perceptions.

The remainder of this article is organized as follows. Section 2 presents the related work on Privacy Engineering and Privacy by Design. Section 3 describes the study settings and survey design. Section 4 reports the empirical findings according to the research questions. Section 5 discusses the implications of the results in light of existing literature. Section 6 outlines the threats to validity, and Section 7 concludes the paper and highlights avenues for future research.

2 Related Work

Privacy Engineering (PE) has sought to integrate data protection into the software lifecycle through methods such as Privacy Impact and Risk Assessments (PIRA), Model based Development (MbD), and Privacy Requirements Engineering (PRE) (Pallas et al., 2024; Sangaroonsilp et al., 2023; Al-Slais, 2020). Complementary efforts include taxonomies grounded in the LGPD Presidência da República do Brasil (2018), ISO/IEC 29100 for Standardization (ISO), and mining techniques for classifying privacy requirements (Sangaroonsilp et al., 2023; Ferrão et al., 2024; Canedo and Mendes, 2020). Despite these advances, systematic approaches to operationalize Privacy by Design (PbD) remain immature. Most contributions focus on requirements or architectural modeling and provide limited empirical validation. Recent work by Kosenkov et al. (Kosenkov et al., 2026) reinforces this point by showing that practices of GDPR and PbD compliance in industry remain largely ad hoc, with little integration between legal requirements, system specifications, and software architecture. These limitations highlight a lack of consolidated and empirically supported diagnostic instruments to evaluate PbD adoption in academic and administrative systems constitutes one of the gaps addressed in the present study.

Beyond technical methods, behavioral and organizational factors also play an important role. Studies show that adoption is influenced by constructs from the Unified Theory of Acceptance and Use of Technology (UTAUT), including performance expectancy, perceived effort, and social influence (Bu et al., 2020, 2021). Barriers such as limited training, fear of sanctions, cultural resistance, and the perception that privacy increases workload are frequently reported. Misconceptions between security and privacy further hinder effective

implementation (Hadar et al., 2018; Ribak, 2019), which reinforces the importance of socio-technical perspectives in Privacy Engineering (Pallas et al., 2024; Canedo et al., 2020).

Recent empirical studies deepen these insights. Spósito et al. (Spósito et al., 2025b) mapped privacy requirements methods (e.g., PriS, LINDDUN, SQUARE) and found adoption concentrated in academia, with limited industrial validation. Rocha and Canedo (Rocha et al., 2023) identified 18 classes of compliance challenges, including legal ambiguity and lack of standardized tools. Matos et al. (Matos et al., 2025) emphasized how cultural norms and misconceptions lead to superficial PbD adoption in organizations. Andrade et al. (Andrade et al., 2024) mapped PbD principles to 72 Privacy Patterns, providing instruments that help reduce the abstraction gap between principles and practice. Andrade et al. (Andrade et al., 2023) also reported that privacy tends to be addressed only in later stages of development, usually without specialists or dedicated tools, which illustrates the immaturity of current practices.

In summary, the literature reveals five recurring gaps: (i) absence of systematic and empirically validated methods for operationalizing PbD, (ii) persistent difficulties in translating legal requirements into concrete engineering activities, (iii) behavioral and organizational barriers that limit adoption, (iv) lack of diagnostic instruments to assess privacy maturity, and (v) reactive treatment of privacy and conceptual conflation between privacy and security. This study builds on these findings by empirically assessing PbD maturity within Brazilian IFES.

3 Study Design

This study aims to investigate the perceptions of Information Technology (IT) professionals from Brazilian 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 Brazilian IFES. The target population consists of IT professionals working in software-related activities across approximately 110 Brazilian IFES. Given the centralized structure of software development teams in these institutions, the number of professionals effectively involved in software development is significantly smaller than the total IT workforce, which constrains the accessible population. 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 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, as well as to coordinators of units responsible for software development within these institutions. Some institutions do not make developer contact information available through their official channels.

Prioritizing, when possible, professionals whose roles were associated with software development or application maintenance. Additionally, the survey employed a screening approach based on explicit participant instructions and characterization questions. The invitation clearly specified the target audience, encouraging self-selection by eligible participants. Furthermore, questions related to professional role (e.g., Q5), experience in IT/software development roles (Q7), and experience with privacy-related tasks (Q9 and Q10) were used to verify that respondents had relevant exposure to software development or related activities. This approach enabled post-hoc validation of participant profiles, although it does not ensure strict exclusion through automated filtering.

The survey questions themselves were carefully designed with direct references to the target group. Invitations were further reinforced through direct messages to practitioners' profiles on social media platforms and, when available, via WhatsApp. Participants were also encouraged to share the survey with colleagues, broadening the outreach within the intended population.

In the context of IFES, IT professionals frequently perform multiple roles, including software development, maintenance, and system evolution activities. These activities are often centralized in specific units, and professionals may alternate between development and maintenance tasks depending on institutional demands. For this reason, the study targeted professionals involved in software-related activities, including both development and application maintenance, as both involve handling personal data and making decisions relevant to Privacy by Design principles. Although the study does not explicitly quantify the proportion of time dedicated to each type of activity, both software development and evolutionary maintenance tasks require decisions that directly impact privacy practices.

The sampling strategy can be characterized as non-probabilistic and based on voluntary participation, which is common in software engineering surveys due to the difficulty of accessing complete population lists. As a consequence, the study prioritizes contextual representativeness and diversity of participants over statistical generalization.

Although the study did not explicitly measure the frequency with which participants engage in software development activities, questions related to professional roles, experience, and involvement in software lifecycle tasks (e.g., Q5 and Q7 - Table 1) were used to ensure that respondents had relevant exposure to software development practices.

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 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 experience with data privacy. All items were closed-ended, except for one open-ended question (Q10) that invited participants to describe their previous experience with privacy.

2. **Privacy Awareness – Knowledge (Q11–Q21):** This section focused on participants' conceptual understanding and knowledge of privacy. It included items assessing familiarity with privacy laws and standards (e.g., LGPD, ISO/IEC 29100, PbD), training experiences, self-learning practices, and recognition of privacy-related risks. One open-ended item (Q13) asked respondents to list five words that come to mind when thinking about privacy, while the remaining questions used closed-ended Likert scales.

3. **Privacy Awareness – Attitudes and Sentiments (Q21–Q27):** This block measured personal beliefs and attitudes regarding privacy, such as distinguishing security from privacy, perceptions of monitoring, relevance of informed consent, and the view of privacy as a fundamental right. Most items were closed-ended Likert scales, with an optional open-ended question for justifying strong disagreements.

4. **Privacy Behaviors and Actions (Q28–Q37):** This section examined concrete actions taken in professional contexts, such as recognizing data retention limits, identifying privacy issues, proposing solutions, and advocating privacy practices within teams. These items were mainly closed-ended Likert scales, complemented by an optional open-ended question to explain disagreements.

5. **Privacy Strategies and Practices (Q38–Q43):** This part investigated the frequency, importance, and perceived ease of application of privacy strategies and techniques (e.g., encryption, minimization, anonymization, risk management, code reviews). Questions were structured as matrix items using Likert-type scales. Two open-ended items (Q40 and Q42) allowed participants to suggest additional strategies or justify their assessments.

6. **Organizational and Individual Factors (Q44–Q55):** This final section addressed organizational dynamics influencing privacy adoption, including the existence of dedicated privacy teams, use of specialized tools, prioritization of privacy, productivity impacts, incentives, sanctions, and perceptions of Privacy by Design. Most items were closed-ended (Likert scales or categorical options), with open-ended questions (Q45 and Q48) inviting respondents to describe their organization's privacy context and future challenges.

In total, the survey combined 46 closed-ended and 9 open-ended questions. Closed-ended items predominantly used five-point Likert scales or multiple choice options, ensuring comparability across responses, while the open-ended items provided deeper qualitative insights into participants' perceptions, experiences, and contextual practices, as shown in Table 1. The full survey instrument is provided as supplementary material available on Zenodo, along with the complete questionnaire and anonymized responses: `https://zenodo.org/records/17247244`.

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 via email.

3.3 Pilot Study

A pilot test was conducted to evaluate the clarity, consistency, and usability of the questionnaire. The form was sent to three

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
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-ended	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

IT staff members from a Federal Institute (IF), whose feedback informed important refinements. They suggested adjustments to the wording of several questions, the removal of redundant items, and the inclusion or modification of response options. Based on their feedback, the questionnaire was revised. Pilot participants took approximately 15-20 minutes to complete the survey, and this average time was communicated to respondents when the final version was released. Pilot responses were not included in the analysis.

3.4 Data Analysis

This study adopts a mixed-method approach, integrating both quantitative and qualitative techniques. To characterize the sample and classify the most cited items, descriptive statistics (e.g., percentages and graphs) were used. For the open-ended questions, we conducted open and axial coding following the principles of Grounded Theory (Bryant and Charmaz, 2007). Grounded Theory is an inductive method for generat-

ing theory directly from data and can be applied to unstructured or structured text, diagrams, images, and even quantitative records (Bryant and Charmaz, 2007).

The coding process involved three stages. First, two authors independently performed open coding of the discursive responses, segmenting the data into discrete parts and generating initial labels. Second, the same authors conducted axial coding, reviewing the segmented data and assigning the codes. Third, categories and subcategories were refined and consolidated. The coding process incorporates respondents' comments, the quotations that supported categories, and the categories and subcategories derived from the data.

To ensure the reliability of the coding process, the two authors initially performed the coding independently. The resulting codes were then compared, and any disagreements were discussed collaboratively until consensus was reached. In cases where discrepancies persisted, a third author reviewed the coding decisions and contributed to the final resolution. This iterative process of discussion and validation helped improve the consistency and robustness of the qualitative analysis.

Although inter-coder agreement was not quantified using statistical measures such as Cohen's kappa, the study adopted a consensus-based approach, which is commonly used in qualitative research following Grounded Theory (Bryant and Charmaz, 2007). This approach prioritizes interpretive alignment among researchers through iterative discussion and refinement of coding decisions, ensuring analytical consistency and reliability.

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

3.5 Inferential Statistical Analysis

To strengthen the quantitative interpretation of the survey data, we complemented the descriptive analysis with inferential statistical procedures. Since most closed-ended questions were measured using Likert-type scales, the data were treated as ordinal. Therefore, we adopted Spearman's rank correlation coefficient (ρ) to examine monotonic associations between selected variables. Spearman's ρ is appropriate for ordinal data and does not assume normal distribution, making it suitable for analyzing relationships among Likert-scale responses.

The inferential analysis was guided by a set of analytical propositions, which are treated in this study as exploratory hypotheses derived from the research questions and theoretical expectations. Rather than testing causal hypotheses, these propositions were used to explore statistically whether perceived importance, ease of use, organizational priority, peer influence, and productivity concerns are associated with the reported use of privacy practices. We defined the following analytical propositions:

- **AP1 – Importance-use alignment:** Privacy strategies perceived as more important tend to be reported as more frequently used.
- **AP2 – Ease-use relationship:** Privacy strategies perceived as easier to apply tend to be reported as more frequently used.
- **AP3 – Organizational priority relationship:** Higher perceived organizational priority regarding privacy and data protection is associated with greater reported adoption of Privacy by Design practices.
- **AP4 – Social influence relationship:** Stronger perceived peer or leadership expectations to adopt privacy practices are associated with greater reported adoption of such practices.
- **AP5 – Productivity concern relationship:** The perception that privacy requirements harm productivity is associated with lower reported adoption of privacy practices.

For each proposition, we selected variables from the questionnaire according to their conceptual relationship with the research questions. Questions Q39, Q41, and Q43 were used to compare the reported use, perceived importance, and perceived ease of privacy strategies. Questions Q47 and Q49–Q55 were used to explore organizational and individual factors related to the adoption of Privacy by Design practices.

For matrix questions involving multiple privacy strategies or development activities, composite indicators were computed by averaging the ordinal scores across the corresponding items. Composite indicators were computed by averaging the corresponding Likert items. Treating these aggregated scores as continuous variables is a widely accepted approximation in empirical software engineering (Norman, 2010).

These composite indicators were used only as exploratory summaries of related Likert-scale items and were interpreted cautiously. The following indicators were considered: a) **Use of privacy strategies** – average score from Q39; b) **Importance of privacy strategies** – average score from Q41; c) **Ease of use of privacy strategies** – average score from Q43; and d) **Use of privacy techniques across software lifecycle activities** – average score from Q38.

Spearman correlation tests were conducted using a significance level of $\alpha = 0.05$. The strength of associations was interpreted as weak, moderate, or strong based on the magnitude of ρ, while the p-value was used to assess statistical significance. The results of these analyses are summarized in Table 7.

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.

The open-ended responses (Q10) reinforced this pattern and further detailed the nature of experience with privacy in software. Many IT professionals explicitly reported having no prior experience, 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) 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.

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%). 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.

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,

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

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

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.

The results presented in Figure 1 (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 Regula-

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

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

tions (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 legislation 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 important foundation of trust between users and developers (Q27), with 43.9% strongly agreeing and 38.6% agreeing, as shown in Figure 1.

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 security-oriented 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 2). 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

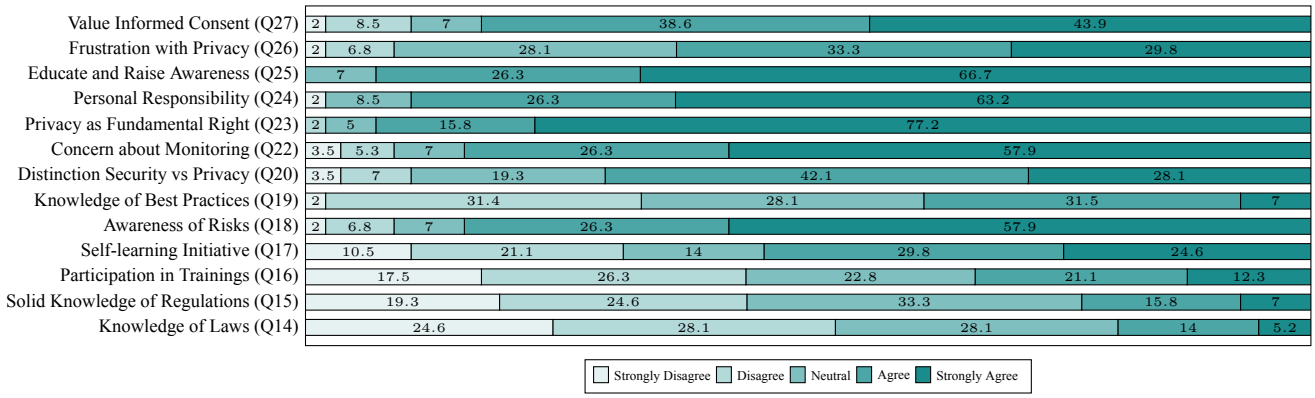

Figure 1. Knowledge, awareness, and attitudes toward privacy (Q14–Q27, $n = 58$).

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 2. 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.

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 3). 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. 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 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 mod-

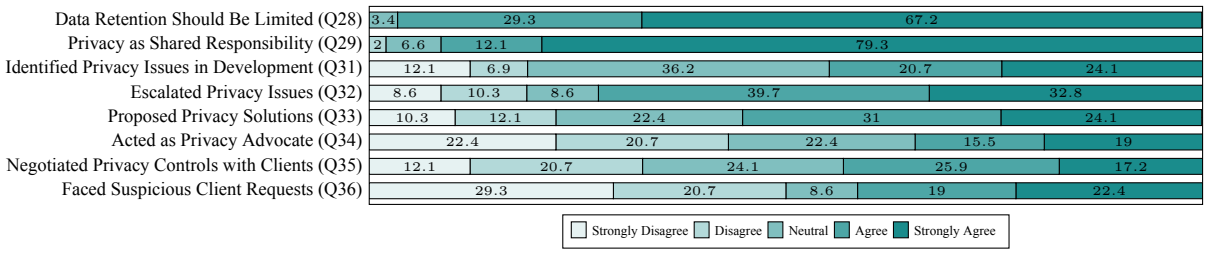

Figure 2. Perceptions about data retention, shared responsibility (Q28–Q29), and practices/experiences related to privacy in software development (Q31–Q36).

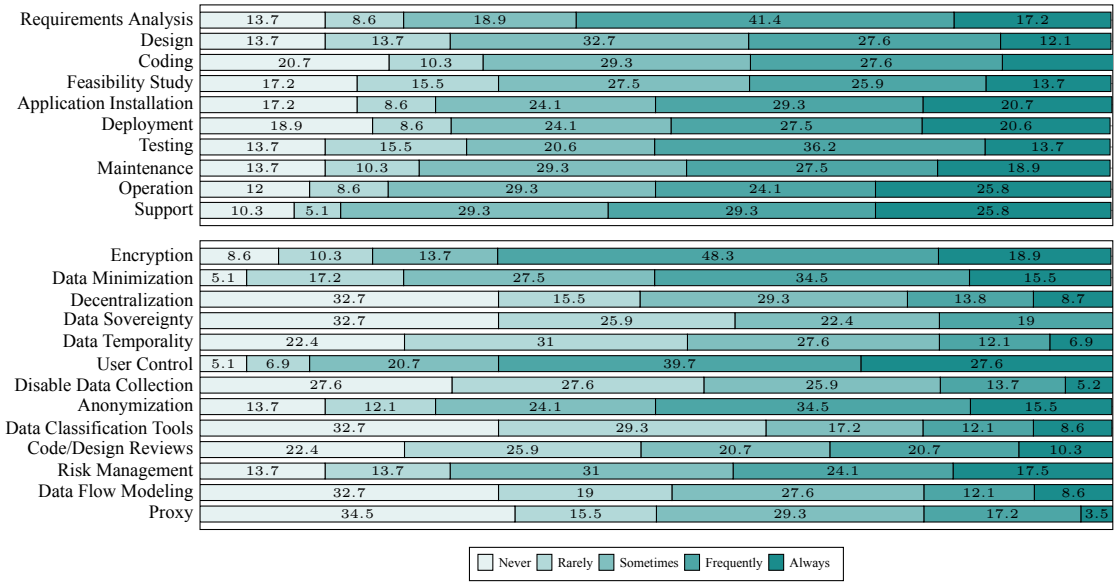

Figure 3. Frequency of using privacy techniques across development activities (top) and specific privacy strategies (bottom), $n = 58$.

erate 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 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.

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.

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, as shown in Table 6. 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(Presidência da República do Brasil, 2018) and GDPR (Union, 2018), 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.

In relation to perceptions of privacy practices (Q49–Q55), the results reveal a heterogeneous picture (Figure 4). 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 4).

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 neutral-

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."

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"

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

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

ity 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 4).

Figure 4. 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, scarce resources, and limited institutional incentives hinder maturity. Although some respondents perceive peer or leadership expectations to adopt privacy practices, the inferential analysis suggests that such social influence is not significantly associated with reported adoption.

4.5 Inferential Analysis of Factors Influencing PbD Adoption

To complement the descriptive results, we conducted inferential statistical tests to examine the relationships between selected variables associated with Privacy by Design adoption. The correlation analysis revealed a statistically significant association only for AP1. The relationship between perceived importance and the reported use of privacy strategies was moderate and statistically significant ($\rho = 0.37$, $p = 0.005$). This result indicates that strategies considered more important by respondents tend to be more frequently adopted in practice.

For AP2, the relationship between perceived ease of use and reported use of privacy strategies was weak and not statistically significant ($\rho = 0.05$, $p = 0.74$), suggesting that perceived usability does not strongly influence adoption. For AP3, the results showed a weak and non-significant association between perceived organizational priority and the use of privacy techniques across the software lifecycle ($\rho = 0.10$, $p = 0.48$).

For AP4, the relationship between social influence and reported use of privacy practices was also weak and not statistically significant ($\rho = 0.06$, $p = 0.64$), indicating that peer or leadership expectations do not appear to be a strong driver of adoption. Finally, AP5 indicated a weak negative association between perceived productivity impact and the use of privacy techniques, which was not statistically significant ($\rho = -0.07$, $p = 0.59$).

Overall, these results indicate that the adoption of Privacy by Design practices in the investigated context is more strongly associated with individual perceptions of importance than with organizational, social, or usability-related factors.

5 Discussion

The findings of this study suggest that, although IT professionals in Brazilian Federal Higher Education Institutions (IFES) demonstrate awareness of privacy as a fundamental right and recognize it as a shared responsibility, the adoption of systematic Privacy by Design (PbD) practices remains limited and largely unsystematic. Furthermore, inferential results indicate that such adoption is more strongly associated with individual perceptions of importance than with organizational or contextual factors.

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., 2025b). 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 indications of this limitation, showing that even within IFES, structured training is rare and practical knowledge about privacy frameworks is insufficient. In this sense, our results also dialogue with the competency oriented framework proposed by Spósito et al. (Spósito et al., 2025a), who systematize privacy and information security competencies for public sector roles aiming to strengthen software quality and LGPD compliance. While their work outlines which knowledge, skills, and attitudes should be present in public sector teams, our data reveal that many IFES professionals still lack these competencies in practice, particularly regarding formal training, regulatory literacy, and familiarity with PbD oriented techniques.

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 provide empirical insights into how these challenges manifest in public sector software development. When interpreted together with the competency framework for public sector roles (Spósito et al., 2025a), these findings reinforce that strengthening PbD maturity requires not only legal interpretation and technical methods, but also clear role definitions and capacity building strategies embedded in institutional governance.

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 con-

Table 7. Spearman correlation results for analytical propositions.

Proposition	Variable 1	Variable 2	ρ	p-value	Interpretation
AP1	Importance (Q41)	Use of strategies (Q39)	0.37	0.005	Moderate, significant
AP2	Ease (Q43)	Use of strategies (Q39)	0.05	0.741	Not significant
AP3	Organizational priority (Q47)	Use across lifecycle (Q38)	0.10	0.480	Not significant
AP4	Social influence (Q53)	Use across lifecycle (Q38)	0.06	0.636	Not significant
AP5	Productivity perception (Q49)	Use across lifecycle (Q38)	-0.07	0.592	Not significant

ceptual 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. The perception that privacy is moderately prioritized by most institutions, coupled with the lack of incentives and recognition for adopting privacy practices, suggests that existing governance structures and leadership signals are insufficient to foster a proactive privacy culture.

In addition to these descriptive findings, the inferential analysis provides further insights into the factors associated with the adoption of Privacy by Design practices. The results indicate that perceived importance is the only factor significantly associated with the reported use of privacy strategies. In contrast, factors such as perceived ease of use, organizational priority, and social influence did not show statistically significant associations with adoption.

This finding suggests that, in the context of IFES, the adoption of privacy practices appears to be driven more by individual perceptions and values than by organizational structures or social pressures. While prior studies emphasize the role of organizational culture and governance in shaping privacy practices, our results indicate that these factors may not translate directly into practical adoption at the team level.

This apparent disconnect between organizational signals and individual behavior reinforces the socio-technical nature of Privacy by Design, highlighting that awareness alone is not sufficient to ensure adoption. Instead, individual prioritization of privacy may act as a key driver, even in environments where institutional support is limited or inconsistent. Although organizational and contextual factors were not statistically significant in the correlation analysis, they remain relevant from a qualitative and interpretive perspective, as reflected in participants' responses and supported by prior literature.

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. At the same time, the low visibility and limited adoption of specialized tools for managing personal data (e.g., privacy management platforms, DLP solutions) suggest that technical infrastructure for PbD remains incipient, which is consistent with the low technological maturity described both in prior case studies (Andrade et al., 2023) and in the competency gaps identified for public sector roles (Spósito et al., 2025a).

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, suggesting that privacy is often treated reactively and often subordinated to other institutional priorities. When juxtaposed with the structured competency framework for privacy and information security in the public sector (Spósito et al., 2025a), our findings suggest a misalignment between the roles and competencies that should exist (according to normative and design oriented proposals) and the current reality of IT teams in IFES, where responsibilities are diffuse, dedicated teams are rare, and training opportunities are scarce.

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 and methodological instruments, but also competency based governance arrangements that clarify responsibilities, support structured training, and align incentives for practitioners in public sector software development. Integrating practitioner perceptions (as captured in our survey) with competency frameworks for privacy and information security (Spósito et al., 2025a) offers a promising path for designing targeted interventions to increase Privacy by Design maturity in Brazilian Federal Higher Education Institutions. In particular, the results highlight that increasing awareness alone may not be sufficient, as adoption appears to depend strongly on how individuals prioritize privacy in their daily practices. Therefore, future interventions should not only focus on organizational policies and tools, but also on reinforcing the perceived importance of privacy among practitioners.

It is important to note that these findings should be interpreted in light of the institutional and organizational context of IFES, including structural, cultural, and economic factors that may influence participants' perceptions and evaluation baselines. As discussed in the Threats to Validity section, these factors may affect how respondents assess the adequacy and maturity of privacy practices. Taken together, these findings provide empirical evidence of the gap between normative frameworks and practical implementation of Privacy by Design in public sector software development, particularly in resource-constrained institutional environments.

6 Threats to Validity

Following Wohlin et al. (Wohlin et al., 2012), the threats to validity in this study are discussed across four dimensions: construct, internal, conclusion, and external validity. Mitigation strategies adopted during the research process are also presented.

Construct validity concerns whether 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 designed based on established frameworks (e.g., LGPD, ISO/IEC 29100, and PbD principles) and refined through a pilot study with three IT professionals from Federal Educational Institutions. Their feedback enabled adjustments to wording, removal of redundancies, and clarification of items to ensure alignment with the research objectives. The combination of closed-ended Likert scale questions and open-ended items helped capture both quantitative and qualitative insights, contributing to broader construct coverage.

Internal validity refers to potential biases or confounding factors that could influence participants' responses. As this research uses a survey method, common threats include misinterpretation of questions, respondent fatigue, and social desirability bias. These risks were mitigated by guaranteeing participant anonymity, which encourages honest responses, and by designing the instrument to be completed in approximately 15-20 minutes, reducing fatigue. Participation was voluntary, minimizing coercion, and screening and characterization questions helped verify that respondents belonged to the intended target audience of IT professionals from Brazilian Federal Higher Education Institutions (IFES).

Given the number of questions in the survey, participant fatigue may have influenced response quality, particularly in later sections of the questionnaire. In total, the survey included 46 closed-ended and 9 open-ended questions. Although the predominance of closed-ended items reduces cognitive effort, it is possible that respondents exhibited reduced attention or engagement over time. This may introduce order effects, where responses to earlier questions are more carefully considered than those provided later. Additionally, the questionnaire did not employ randomization of question blocks, which could have mitigated such effects. Therefore, responses to later items should be interpreted with caution.

The study did not explicitly measure the proportion of time participants dedicate to software development versus maintenance or support activities. Given the heterogeneous nature of IT roles in IFES, professionals may have varying levels of direct involvement in software development tasks. Although role-related questions (e.g., Q5 and Q7) were used to ensure that respondents had relevant experience in software-related activities, differences in the degree of involvement in development tasks may influence participants' perceptions of Privacy by Design practices, representing a potential source of variability in the responses. In some cases, responses may reflect indirect experience or organizational perceptions rather than recent hands-on involvement in software development activities, particularly for participants working in support, infrastructure, or distributed campus contexts. This may dilute stronger effects and contribute to more moderate aggregated results.

Additionally, the inclusion of professionals with varying degrees of involvement in software development, including those primarily engaged in support or infrastructure activities, may introduce additional variation in the responses. While these participants were not excluded, as the study aims to capture perceptions within the broader organizational context of IFES, their perspectives may differ from those of professionals directly involved in software development. This may influence how certain practices are perceived and reported, particularly in relation to technical implementation aspects of Privacy by Design.

A significant portion of respondents reported having no prior direct experience with privacy in software. This may influence how participants interpret and respond to questions related to Privacy by Design practices, as some answers may be based on general understanding or indirect exposure rather than practical experience. This factor may further contribute to variability in responses and to attenuated effects in the overall results.

The broader organizational context of IFES may introduce additional sources of bias. As public servants, participants operate within an environment characterized by stable employment conditions, reduced exposure to market pressures, and different accountability structures compared to private sector organizations. These factors can shape how respondents perceive technical challenges, prioritize privacy-related issues, and evaluate the urgency of addressing them. This context may contribute to the normalization of structural limitations, such as fragmented governance or delayed adoption of practices.

Budgetary constraints and resource limitations commonly observed in IFES may also influence how participants evaluate the adequacy of privacy practices. In such contexts, respondents may adopt more pragmatic perspectives, assessing existing practices as "sufficient" relative to institutional constraints rather than against externally benchmarked standards. This may lead to a normalization of limitations, where positive evaluations reflect organizational resilience and adaptation rather than objectively high levels of maturity. Consequently, perceptions reported in the study may be conditioned by context-specific evaluation baselines, which should be considered when interpreting and comparing results across different organizational environments.

The timing of the data collection may also have introduced a temporal bias. The survey was conducted between July and September, with part of this period coinciding with vacation schedules in Brazilian Federal Higher Education Institutions. This may have affected the availability of some participants and influenced response patterns. Although the extended collection period helped mitigate this effect, it cannot be fully excluded.

Finally, institutional desirability bias may also be present, as respondents are both users and, in some cases, developers of the systems that process their own personal and professional data. As a result, responses may be influenced by personal concerns regarding data exposure or by a tendency to present institutional practices in a more favorable manner. This may particularly affect responses related to perceived importance of privacy, trust in institutional practices,

and self-reported adoption of Privacy by Design principles. In this sense, although participants were informed that their responses would be treated anonymously, without individual or institutional identification, this procedure helps reduce this type of bias, even though its complete elimination cannot be ensured.

The use of generative AI tools by respondents during survey completion cannot be fully ruled out. While the predominance of closed-ended and perception-based questions reduces the likelihood of substantial external assistance, the possibility of AI-supported responses may affect the authenticity of some answers, particularly in open-ended items. This represents a potential, although limited, source of bias in self-reported data.

Conclusion validity relates to whether the analysis supports sound and reliable conclusions. To enhance this validity, the study employed primarily descriptive statistical analyses, complemented by exploratory correlation analysis and qualitative coding using Grounded Theory for open-ended responses. While these analyses support the identification of patterns and relationships within the data, the exploratory nature of the correlation tests and the limited sample size constrain the strength of statistical generalizations. To mitigate this, quantitative findings were triangulated with qualitative insights, increasing the robustness and interpretive depth of the results.

External validity concerns the extent to which findings may be generalized beyond the study sample. The survey included 58 IT professionals across 15 Brazilian states, providing geographic and institutional diversity. However, this sample represents a restricted subset of IT professionals, as the study specifically targets individuals involved in software-related activities within IFES, where such roles are typically centralized in dedicated units.

We acknowledge that the sample size may limit statistical generalization. However, in the context of empirical software engineering, obtaining large and fully representative samples is inherently challenging due to the lack of comprehensive population frames and the voluntary nature of participation (Kitchenham and Pfleeger, 2002, 2008; de Mello and Travassos, 2016). As highlighted in prior work, representativeness depends not only on sample size but also on the adequacy of the target population definition and the diversity of the sample.

To mitigate this limitation, the study ensured diversity across institution types (Federal Universities and Federal Institutes) and geographic regions. Nevertheless, the findings should be interpreted as indicative of trends within the investigated context rather than statistically generalizable to the entire population of IFES professionals. The results were also compared with findings from related studies (e.g., (Matos et al., 2025; Rocha et al., 2023)) to strengthen external consistency.

7 Conclusion

This study investigated how Privacy by Design (PbD) is perceived and practiced within Brazilian Federal Higher Education Institutions (IFES) through a mixed-method survey with 58 IT professionals from 15 states. The findings indicate that the level of maturity in the implementation of Privacy by Design principles within Brazilian Federal Higher Education Institutions can be characterized as emerging, fragmented, and uneven across institutions and teams, reflecting heterogeneous levels of adoption and organizational support.

The findings reveal strong recognition of privacy as a fundamental right and a shared responsibility, as well as frequent use of core strategies such as encryption, minimization, and anonymization. Nevertheless, substantial gaps remain in regulatory knowledge, structured training, and the systematic adoption of more advanced privacy-enhancing practices. The inferential analysis further revealed that the adoption of Privacy by Design practices is significantly associated with individual perceptions of importance, while organizational, social, and usability-related factors do not show statistically significant relationships with adoption. This suggests that, in the investigated context, privacy practices are driven more by individual prioritization than by institutional structures or external pressures.

While there is evidence of awareness and partial adoption of privacy practices, the results reveal gaps in systematic implementation, organizational support, and integration of privacy throughout the software development lifecycle.

These results complement and extend previous research. They confirm the academic practice gap identified by Spósito et al. (Spósito et al., 2025b,a), exemplify the compliance challenges reported by Rocha and Canedo (Rocha et al., 2023), and reinforce the socio-technical barriers described by Matos et al. (Matos et al., 2025). In line with the observations of Andrade et al. (Andrade et al., 2024, 2023), our findings underscore the need for practical instruments, empirical validation, and greater operational support for implementing PbD in real-world settings.

Advancing PbD maturity in IFES will require coordinated and sustained efforts that include structured and continuous training, clearer definition of roles and responsibilities, adoption of privacy-supporting tools, and cultural change to embed privacy as a strategic institutional value.

While awareness is increasingly present, maturity remains incipient and uneven. The results indicate that the adoption of Privacy by Design practices is strongly influenced by the perceived importance attributed by practitioners, suggesting that individual prioritization plays a central role in driving adoption. Therefore, improving PbD maturity requires not only organizational policies and tools, but also targeted efforts to reinforce the perceived importance of privacy among IT professionals. Such advancements are essential not only for ensuring LGPD compliance but also for reinforcing trust in academic systems and consolidating privacy engineering practices within Brazil's public higher education sector.

Future work may further investigate how frequently IT professionals in IFES engage in software development activities, as well as how differences in professional roles, levels of experience, and degrees of involvement in software development influence perceptions and adoption of Privacy by Design practices. In particular, subgroup analyses could provide more fine-grained insights into how these factors shape the variability observed in this study.

Artifact Availability

All artifacts produced in this study are openly available on Zenodo at `https://zenodo.org/records/17247244`. The repository includes the full survey instrument, the complete questionnaire with all response options, and the anonymized dataset of participants' answers.

Acknowledgment

This study was financed in part by the National Council for Scientific and Technological Development (CNPq), Grants No. 300883/2025-0 and No. 406266/2025-5.

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.

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.

Canedo, E. D. and Mendes, B. C. (2020). Software requirements classification using machine learning algorithms. *Entropy*, 22(9):1057.

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.

de Mello, R. M. and Travassos, G. H. (2016). Surveys in software engineering: Identifying representative samples. In *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM 2016, Ciudad Real, Spain, September 8-9, 2016*, pages 55:1–55:6. 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.

for Standardization (ISO), I. O. (2024). Iso/iec 29100:2024. information technology — security techniques — privacy framework.

Gonçalves, André Luiz Dias (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`.

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.

Kitchenham, B. A. and Pfleeger, S. L. (2002). Principles of survey research: part 5: populations and samples. *ACM SIGSOFT Softw. Eng. Notes*, 27(5):17–20.

Kitchenham, B. A. and Pfleeger, S. L. (2008). Personal opinion surveys. In Shull, F., Singer, J., and Sjøberg, D. I. K., editors, *Guide to Advanced Empirical Software Engineering*, pages 63–92. Springer.

Kosenkov, O., Zabardast, E., Fucci, D., Méndez, D., and Unterkalmsteiner, M. (2026). Privacy by design: Aligning GDPR and software engineering specifications with a requirements engineering approach. *Inf. Softw. Technol.*, 190:107946.

Kuliamin, V. V., Petrenko, A. K., and Rudina, E. A. (2025). Software security by design. *Program. Comput. Softw.*,

51(6):429–434.

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).

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, 15(5):625–632.

Pallas, F., Koerner, K., Barberá, I., Hoepman, J., Jensen, M., Narla, N. R., Samarin, N., Ulbricht, M., Wagner, I., Wuyts, K., and Zimmermann, C. (2024). Privacy engineering from principles to practice: A roadmap. *IEEE Secur. Priv.*, 22(2):86–92.

Presidência da República do Brasil (2018). Lei nº 13.709, de 14 de agosto de 2018. lei geral de proteção de dados pessoais (LGPD). `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. and Canedo, E. D. (2025). Optimizing compliance: Comparative study of data laws and privacy frameworks. *J. Internet Serv. Appl.*, 16(1):431–452.

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., Alves, K., Nunes, R., Ferreira, L., and Canedo, E. (2025a). Structuring privacy and information security competencies for public sector roles: A framework for enhancing software quality and lgpd compliance. In *Anais do XXIV Simpósio Brasileiro de Qualidade de Software*, pages 365–375, Porto Alegre, RS, Brasil. SBC.

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. (2025b). 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-pelo-governo`.

Union, E. (2018). General data protection regulation (GDPR). *Intersoft Consulting*, 1(1):1–100.

United Nations (ONU) (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-12`.

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