Assuring Trustworthy Data: A Dual-Criteria Analysis of Anonymization and System Reliability in Digital Health (A Systematic Review)

Giovana Nunes Inocêncio[1], **Jean Everson Martina**[1]

[1] Universidade Federal de Santa Catarina (UFSC)

gioinocencio017@gmail.com, jean.martina@ufsc.br

Abstract. ***Research Context:*** *The Digital Health sector faces the critical challenge of reconciling clinical data privacy with system reliability and utility. The risk in Brazil is high; the sector is the most targeted by ransomware, facing catastrophic financial losses.* ***Scientific and/or Practical Problem:*** *A persistent gap exists in the literature concerning the systematic evaluation of how anonymization techniques impact integrity and the lack of consensus on the dual criteria (technical and sociotechnical) necessary to assess effectiveness in Health Information Systems (HIS).* ***Proposed Solution and/or Analysis:*** *A Systematic Literature Review (SLR) mapped anonymization techniques (RQ1), analyzed their reliability impact (RQ1.1), and identified formal criteria for user trust (RQ2), addressing the core privacy-utility dilemma.* ***Related IS Theory:*** *The study is based on the Sociotechnical Systems Theory (SST), recognizing that anonymization and reliability are outcomes of integrating technology with critical social factors, including governance and user trust. The research aligns with GranDSI-BR Challenges 2 and 4.* ***Research Method:*** *The SLR followed Kitchenham and PRISMA guidelines, utilizing the PICOC model. The search across four major databases yielded 20 high-quality articles published between 2020 and 2025.* ***Summary of Results:*** *Dominant techniques are Blockchain and Federated Learning (FL), substantially enhancing data integrity and the privacy-utility balance. Reliability is dually assessed by Technical criteria (e.g., Re-identification Risk) and Sociotechnical criteria (e.g., Governance, Public Perception).* ***Contributions and Impact to IS area:*** *The study consolidates a theoretical and empirical framework on anonymization's influence on reliability, meeting GranDSI-BR Challenges 2 and 4. It offers practical subsidies for managers and regulators in designing verifiable and trustworthy digital health systems.*

1. Introduction

The advancement of digital health and the use of sensitive clinical data have intensified the challenge of balancing patient privacy with the reliability and usability of healthcare systems. This dilemma is particularly critical in Brazil, where the healthcare sector is the most targeted by cyberattacks, including a 69% growth in targeted ransomware activity [Kaspersky 2025]. The critical nature of patient data makes the sector a primary target, exposing institutions to significant operational and financial liabilities. Beyond the high frequency of attacks, security failures pose a catastrophic financial risk: the healthcare sector incurs the highest cost per data breach incident in the country, resulting in losses of R$ 11.43 million [Ponemon Institute and IBM 2025]. Given this elevated risk, data anonymization emerges as a fundamental tool for the ethical and secure use of clinical information.

In this context, anonymization is a crucial data protection method that seeks to eliminate or modify elements that would otherwise allow for the direct or indirect identification of data subjects, thereby ensuring the privacy and confidentiality of the information [Sposito et al. 2024]. The core consequence of a lack of anonymization is a failure to meet essential data protection standards, which inherently exposes sensitive records to the constant threat of unauthorized access, accidental destruction, loss, alteration, or any form of inappropriate handling [Sposito et al. 2024]. Reliability is a key factor in the adoption of Health Information Systems (HIS) and depends on the interplay between human, process, and technological elements. At the relational level (People/Processes), it relies on users' willingness to share data, grounded in trust. This trust, in turn, is tied to system reliability, which at the technological level (Processes/Technology) demands strong anonymization mechanisms.

Anonymization is crucial for both legal compliance and ethical use; however, a persistent gap exists in the literature regarding a systematic evaluation of how different techniques impact system reliability and integrity. Furthermore, there's a lack of consensus on the dual criteria (technical and sociotechnical) needed to assess de-identification effectiveness. This shortfall impedes managers and regulators from confidently designing trustworthy digital health solutions. Therefore, this study has practical and social relevance, providing insights for healthcare managers and regulatory authorities to enable more informed decisions on clinical data protection and digital healthcare governance.

This study addresses this critical need by offering the first systematic assessment that proves the effectiveness of anonymization is dually determined by both technical robustness and sociotechnical criteria. Our research is novel in that it consolidates an empirical framework that maps the influence of specific technologies (Blockchain, FL) on technical trustworthiness, while simultaneously establishing the non-algorithmic requirements essential for preserving patient trust (RQ2). The study's theoretical contribution is to consolidate a systematic review of the literature that seeks to answer the question: **'How do clinical data anonymization techniques influence the reliability of digital health systems?'**.

The approach of this study aligns directly with Sociotechnical Systems Theory, treating reliability as an outcome of integrating these dual factors. This systemic complexity is examined through the lens of Sociotechnical Systems Theory, recognizing that reliability is an outcome of integrating technology (such as anonymization techniques) with social factors (including governance and user trust) [Mumford 2006]. This contribution directly aligns with the GranDSI-BR Challenges (specifically Challenge 2 on Open World and Challenge 4 on Sociotechnical Vision), addressing the critical need to ensure governance and trustworthy use of sensitive health data in Brazil's interconnected digital transformation.

The article is organized as follows: Section 2 presents the theoretical background on anonymization and reliability; Section 3 details the Systematic Literature Review

(SLR) methodology, including the protocol and quality criteria; Section 4 presents the results and analysis of the techniques and criteria mapped; and Section 5 discusses the study's conclusions and contributions.

2. Background

The rapid digitalization of clinical data creates an inherent tension between protecting sensitive patient information and ensuring the data's utility for analysis [Queiroz et al. 2016]. This necessity defines anonymization as the core challenge, requiring the use of available and reasonable technical means to prevent direct or indirect association with a person. Systematic privacy preservation methods classify data (e.g., Explicit Identifiers, Quasi Identifiers, and Sensitive Attributes) [Carvalho et al. 2021] and employ operations like Suppression and Generalization to satisfy formal Privacy Models (e.g., k-Anonymity or ℓ-Diversity Principle) [Carvalho et al. 2021]. However, these modifications lead to a quantifiable loss of data quality, creating the central Privacy and Utility trade-off, often measured by Information Metrics such as ILoss [Camêlo and Alves 2023].

In Health Information Systems (HIS), reliability is a complex, dualistic concept viewed through the lens of Sociotechnical Systems Theory [Camêlo and Alves 2023]. Technical reliability demands mandated Security measures to protect against data loss or alteration. Crucially, sociotechnical reliability incorporates User Trust and ethical governance, ensuring compliance with ethical principles, such as non-discrimination [Carvalho et al. 2021]. The literature confirms that translating these legal and ethical requirements into functional systems is a significant challenge for IT professionals, who often lack sufficient knowledge of legislation such as the Lei Geral de Proteção de Dados (LGPD) [Camêlo and Alves 2023].

The current State of the Art emphasizes structured approaches, including domain ontologies, to unify terms and provide the ”reasonable technical means” required by the LGPD [Queiroz et al. 2016]. While the LGPD is the legal foundation in Brazil, mandating requirements like Privacy by Design (PbD), its abstract legal text creates a significant sociotechnical hurdle for operationalizing privacy requirements [Camêlo and Alves 2023].

A fundamental gap persists: Although the Privacy-Utility dilemma is known, and both classic (generalization/suppression) and advanced (Blockchain/FL) techniques exist, no previous SLR has performed a systematic and comprehensive evaluation of the direct, quantifiable relationship between specific anonymization techniques and the dual criteria of system reliability in HIS. Research is needed to consolidate how technical choices impact both Technical Risk/Utility (re-identification risk) and Sociotechnical Trust/Acceptance (ethical governance/patient confidence) [Camêlo and Alves 2023]. This SLR directly addresses this deficiency by consolidating an integrated framework for trustworthy and legally compliant digital health systems.

3. Method

To address the critical research gap in balancing data privacy and utility in clinical settings, we conducted a Systematic Literature Review (SLR). This methodology was strategically chosen given the multifaceted nature of the problem, which involves technical, ethical, and legal dimensions, requiring an approach focused on the global scientific context and the intersection of these factors in Health Information Systems. The SLR adhered to the PRISMA guidelines (establishing the standard for reporting and transparency) [Moher et al. 2009] and the Kitchenham Checklist [Kitchenham and Charters 2007]. The PICOC model [Wohlin et al. 2012] guided the formulation of the research questions. To ensure the inclusion of high-quality articles, a methodological quality checklist, based on the Kitchenham approach [Kitchenham and Charters 2007], was applied.

3.1. PICOC and Research Questions

The PICOC model defines five essential elements: the group of interest (P), the technology under study (I), the control treatment (C), the practical outcomes (O), and the research environment (Context) [Wohlin et al. 2012]. This framework is fundamental for providing a structured and systematic framework to define the scope of the SLR, thereby ensuring that the research questions (RQs) are focused, clear, and address a specific area of inquiry, as shown in Table 1.

Table 1. PICOC framework used in the study.

PICOC Element	Study Definition
P (Population)	Digital health systems that use clinical data.
I (Intervention)	Application of Clinical Data Anonymization or De-Identification Techniques.
C (Comparison)	Not applicable. The study focuses on analyzing the implications of anonymization across the literature, rather than conducting an experimental comparison.
O (Outcome)	Reliability of anonymized systems and data, including aspects such as user trust, integrity, quality, and information security.
C (Context)	Digital health environments, including clinical, administrative, research, or health information governance applications.

The research questions defined for this SLR were:

- **RQ1:** What clinical data anonymization techniques are used in digital health systems?
  - **RQ1.1:** How do these techniques influence the reliability of these systems?
- **RQ2:** What criteria are used to evaluate the effectiveness of anonymization in preserving user trust?

The Intervention (I) and Context (C) defined RQ1 by focusing on identifying specific anonymization techniques in digital health. The Outcome (O) of Reliability drove RQ1.1, linking these techniques (I) to system trustworthiness. Finally, the reliance of (O) on user trust established RQ2, mapping criteria beyond technical metrics for assessing the effectiveness of anonymization.

3.2. Inclusion, Exclusion, and Selection Criteria

To ensure consistency of the studies included in the SLR, inclusion, exclusion, and selection criteria were defined to guide the screening process, ensuring that only studies aligned with the research objective were analyzed. Table 2 summarizes these criteria. The criteria related to language (IC2) and publication year (IC1) were applied directly in the search databases, acting as the initial filters of the Selection Process (SC5).

Table 2. Systematic Review Inclusion, Exclusion, and Selection Criteria.

Inclusion Criteria (IC)
IC1: Studies published between 2020 and 2025. IC2: Written in English or Portuguese. IC3: Primary studies focused on at least one of the following topics: • Clinical data anonymization techniques. • Assessment of the reliability or integrity of anonymized data. • Applications in digital health systems. • Proposals and evaluations of frameworks, models, or metrics.
Exclusion Criteria (EC)
EC1: Duplicate studies in more than one search database. EC2: Studies that address general safety, without specifically addressing the reliability of anonymized data. EC3: Studies that are not complete, such as abstracts and opinion articles.
Selection Criteria (SC)
SC1: Perform an automatic search in the databases, with the string adapted to the specific syntax of each database. SC2: Remove duplicates using Zotero’s features, with manual checking when necessary. SC3: Read titles and abstracts to identify potentially relevant studies in Parsifal platform. SC4: Apply the inclusion and exclusion criteria defined in this protocol. SC5: Read the full-text of the pre-selected studies. SC6: Assess the methodological quality of the studies, based on the Kitchenham checklist [Kitchenham and Charters 2007].

3.3. Search String and Databases

The search string, as shown in Table 3, was defined based on the main terms related to the search topic, ensuring comprehensive coverage through the inclusion of synonyms and spelling variations (such as *anonymization/anonymisation*). The string encompasses key concepts related to digital health, data anonymization, reliability, and clinical systems. Its structure utilized Boolean operators (AND, OR) to enhance search precision.

Table 3. Search String used for the SLR.

```
("health system*" OR "health information system*" OR "eHealth" OR "digital
health") AND ("anonymization" OR "anonymisation" OR "data anonymization" OR "data
de-identification" OR "privacy-preserving") AND ("trust" OR "trustworthiness" OR
"reliability" OR "data quality" OR "confidence" OR "data integrity") AND ("clinical
data" OR "health data" OR "medical data")
```

The search strategy deliberately excluded modern techniques to ensure methodological breadth and prevent bias, allowing the review to naturally identify which anonymization techniques the literature highlights as most relevant to reliability, thus addressing RQ1 comprehensively. We included the terms ”reliability” and ”trust” to capture the system’s dual sociotechnical dimensions. Reliability, which reflects data integrity and technical assurance, directly aligns with RQ1.1. Trust represents user confidence and ethical governance, aligning with RQ2.

The SLR involved a comprehensive search in databases chosen for their relevance to the areas of digital health, computer science, and engineering. The PubMed search database offers broad interdisciplinary coverage (health and technology), IEEE Xplore and ACM Digital Library are important for more technical studies, and SpringerLink complements this with applied studies. The number of results of the search string in each database is shown in Table 4.

Table 4. Number of results per search base.

Search base	Number of results
IEEE Xplore [1],	14
PubMed [2],	6
ACM Digital Library [3],	255
SpringerLink [4],	1007

In the IEEE database, Boolean operators and truncation with an asterisk (*) were employed. In PubMed, since truncation is not supported, the ”[Title/Abstract]” field was added to limit searches to titles and abstracts. In the ACM Digital Library, the ”[All:]” syntax was applied for full-content searches without truncation. Finally, in SpringerLink, a similar search string to IEEE’s was used, combining Boolean operators and truncation, which enabled a comprehensive search.

3.4. Methodological Quality Criteria

An assessment of methodological quality was conducted to ensure the reliability and relevance of the evidence, utilizing the adapted Kitchenham Checklist [Kitchenham and Charters 2007], as shown in Table 5.

Table 5. Methodological Quality Criteria (Kitchenham Checklist).

Code	Criterion	Weight
QC1	Is the objective of the study clearly presented, with context, problem, and contribution described?	1
QC2	Is the anonymization method described in the steps, tools, and parameters used?	1
QC3	Are there objective metrics or criteria to assess reliability?	1
QC4	Was the usefulness of anonymized data discussed?	1
QC5	Was the study applied to real data or a practical scenario?	1
QC6	Does the study discuss limitations or risks to reliability?	1

The six defined criteria (QC1–QC6) were tailored to filter studies based on their technical rigor and contextual completeness necessary for answering the RQs, focusing on the balance between technical application (QC2, QC5) and the core privacy-utility trade-off (QC3, QC4). Only studies scoring at least four points were included in the review.

[1] IEEE Xplore (`https://ieeexplore.ieee.org`)

[2] PubMed (`https://pubmed.ncbi.nlm.nih.gov`)

[3] The ACM Digital Library (`https://dl.acm.org`)

[4] SpringerLink (`https://link.springer.com`)

3.5. Selection Process

One author conducted the systematic literature review (SLR), and to ensure the rigor of the screening process, a second author independently validated the application of the inclusion and exclusion criteria and the initial data extraction of eligible studies. This independent validation acted as the mechanism for analyzing the rigor and concordance of the protocol.

Studies published between 2020 and 2025 in English or Portuguese were selected using database-specific search strings. Two reviewers independently conducted study selection and in-depth analysis, reaching consensus on disagreements, and extracted evidence on anonymization techniques, their impact on system reliability, and user trust criteria.

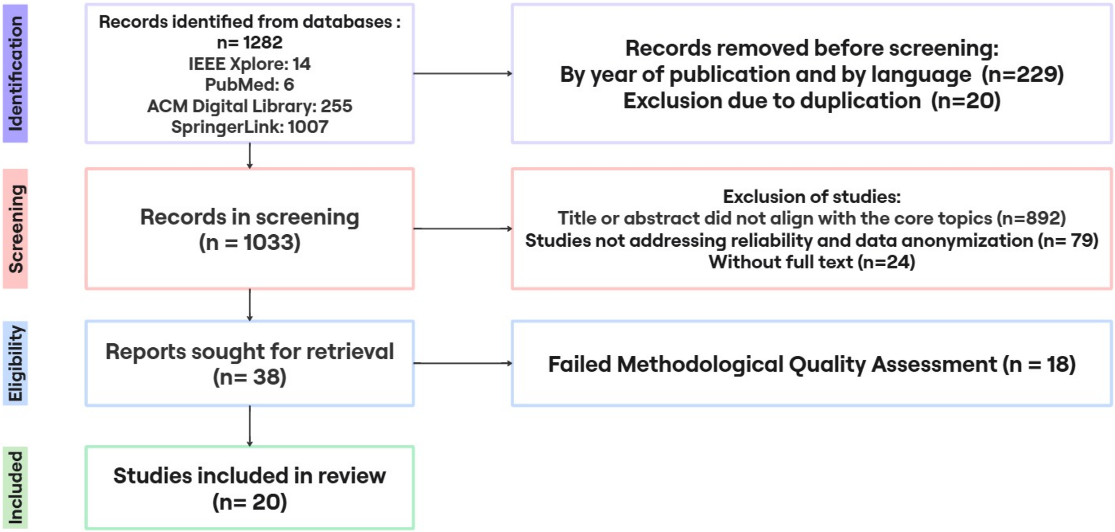
Identification
Records identified from databases : n= 1282
IEEE Xplore: 14
PubMed: 6
ACM Digital Library: 255
SpringerLink: 1007
Records removed before screening:
By year of publication and by language (n=229)
Exclusion due to duplication (n=20)
Screening
Records in screening (n = 1033)
Exclusion of studies:
Title or abstract did not align with the core topics (n=892)
Studies not addressing reliability and data anonymization (n= 79)
Without full text (n=24)
Eligibility
Reports sought for retrieval (n= 38)
Failed Methodological Quality Assessment (n = 18)
Included
Studies included in review (n= 20)

Figure 1. PRISMA flow diagram of the systematic literature review.

The studies underwent stages of identification, screening, eligibility, and inclusion, as illustrated in Figure 1, following the selection criteria (SC) from Table 2. After collecting, removing duplicates, and initially organizing references via the Zotero plugin, the data was exported to the Parsifal tool for managing the final pool and facilitating preliminary filtering.

The study was conducted by applying the strings to the selected search databases, adhering to the inclusion criteria (IC) and exclusion criteria (EC) outlined in the PRISMA flowchart [Moher et al. 2009] (Table 2). The search was conducted in May 2025 using the specified search strings in the selected databases, yielding 1,282 results. Next, the language and year of publication filters were applied directly to the search databases (IC1 and IC2), considering only studies in English or Portuguese published between 2020 and 2025 for inclusion. With this restriction, 229 records that did not meet the time criterion were excluded, resulting in a total of 1,053 references in the search databases. The obtained references were imported into the Parsif.al platform, where the process of identifying and removing duplicate records (EC1) was performed, resulting in 1,033 results.

Subsequently, screening was performed, with a preliminary filtering based on the analysis of titles and abstracts, to select studies compatible with topics related to clinical data anonymization, reliability of anonymized data, applications in digital health systems, and proposed frameworks or metrics (IC3). A total of 141 studies considered potentially relevant were selected based on the title screening, of which 79 studies were excluded because they addressed safety in general terms, without specifically discussing data reliability (EC2). Additionally, 24 studies were rejected because they were opinion articles or abstracts without full-text availability (EC3). This 'grey literature' was excluded as it typically lacks the detailed methodological reporting necessary to provide reliable evidence on anonymization techniques and their impact on system reliability.

Thus, 38 articles remained and were accepted for further analysis. Of these, 33 were selected for meeting all inclusion criteria, such as publication between 2020 and 2025 (IC1), written in English or Portuguese (IC2), and being primary studies focused on clinical data anonymization techniques, assessing the reliability or integrity of anonymized data, applications in digital health systems, and proposed frameworks related to the topic (IC3). The remaining five articles were also accepted because, although they did not meet the criteria for primary studies (IC3), they did not violate the exclusion criteria and presented approaches that can contribute to a deeper understanding of the field. The 38 eligible studies underwent a Methodological Quality Assessment, according to the Kitchenham checklist [Kitchenham and Charters 2007]. Applying the cutoff threshold, 20 articles were included in this study, as detailed in Subsection 3.6.

3.6. Application of Quality Criteria

Based on the 38 eligible articles, a quality assessment of the studies was performed, as shown in Table 6, aiming to meet the six quality criteria established in Table 5.

Table 6. Distribution of Included Studies by Publication Year and Type of Study.

Ref.	Points	Status	QC1	QC2	QC3	QC4	QC5	QC6
[Kumar et al. 2024]	6.0	Incl.	Y	Y	Y	Y	Y	Y
[Vovk et al. 2021]	6.0	Incl.	Y	Y	Y	Y	Y	Y
[Raghav and Bhola 2023]	6.0	Incl.	Y	Y	Y	Y	Y	Y
[Tseng and et al. 2025]	6.0	Incl.	Y	Y	Y	Y	Y	Y
[Kathole et al. 2024]	5.5	Incl.	Y	Y	P	Y	Y	Y
[Chenthara and et al. 2020]	5.5	Incl.	Y	Y	P	Y	Y	Y
[Anusuya and et al. 2022]	5.5	Incl.	Y	Y	Y	Y	P	Y
[Madhavi et al. 2024]	4.5	Incl.	Y	Y	P	Y	P	P
[Wu et al. 2021]	4.5	Incl.	Y	Y	P	Y	P	P
[Khaled and Ali 2025]	4.5	Incl.	Y	Y	P	P	Y	P
[Purohit et al. 2025]	4.5	Incl.	Y	Y	P	P	Y	P
[Monteiro and et al. 2024]	4.5	Incl.	Y	Y	P	Y	P	P
[Sami and Toorani 2024]	5.0	Incl.	Y	Y	P	P	Y	Y
[Fakeeroodeen and Beeharry 2021]	5.0	Incl.	Y	Y	P	Y	Y	P
[Herwanto and et al. 2024]	5.0	Incl.	Y	P	Y	Y	P	Y
[Watkins and et al. 2023]	5.0	Incl.	Y	Y	P	Y	Y	P
[Zala et al. 2024]	4.0	Incl.	Y	Y	P	P	P	P
[Tian and et al. 2024]	4.0	Incl.	Y	N	Y	P	P	Y
[Churi and Pawar 2024]	4.0	Incl.	Y	Y	P	P	P	P
[Dotter and et al. 2025]	4.0	Incl.	Y	N	N	Y	Y	Y

Legend: *Ordered by score DESC.*

- **Points:** Total quality score (Maximum 6.0). Included Studies: Score $\geq$ 4.0.
- **Status:** Incl. (Included), Excl. (Excluded).
- **QC1-QC6:** Quality Criteria (Y: Yes = 1.0, P: Partial = 0.5, N: No = 0).

The methodological quality of the 38 eligible articles was assessed using the six criteria (QC1–QC6) detailed in Table 5. Based on the scoring threshold, 20 articles with a score of 4.0 or higher were included in the review. The remaining 18 articles were excluded due to insufficient quality, primarily related to limitations in the description of the method, reliability assessment, or discussion of constraints, ensuring that the final analysis relies only on high-quality, relevant evidence.

Snowballing was deliberately not adopted in this review due to time and scope constraints, and to keep the protocol focused on a reproducible database-driven search process. Future work can extend this review by applying systematic backward and forward snowballing procedures, as recommended by Wohlin, to complement the database search and broaden the evidence base[Wohlin et al. 2012].

4. Results and Discussion

4.1. Overview and Characterization of Included Studies

The results of this study are based on 20 high-quality articles that followed the protocol outlined in Section 3. Analyzing the number of studies per year, we can see that in recent years the number of studies aligned with this study's protocol has increased, particularly in 2024 and 2025, as shown in Figure 2. While one journal article was identified in 2020, the number of results in 2021 increased to four studies (two journal articles and two conference proceedings). In 2022 and 2023, the number of conference proceedings was reduced to just one per year. Most studies were published in 2024 (nine) and 2025 (four), accounting for 65% of the literature. This concentration indicates recent growth and sustained research interest in anonymization and reliability techniques for digital health systems.

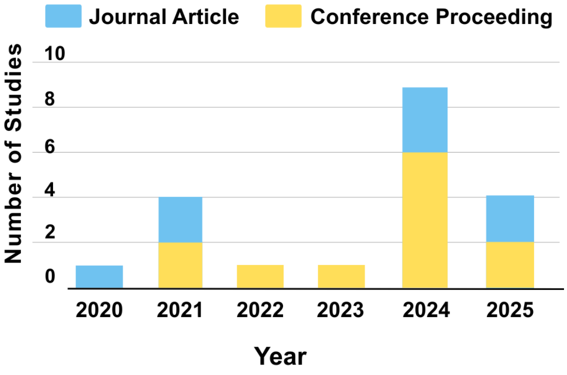

Figure 2. Distribution of Included Studies by Publication Year and Type of Study.

The distribution analysis of the 20 included studies reveals a focus on generating new technical artifacts. There were 55% (11 studies) about Design and Solution Development. This data highlights that the majority of research efforts focus on proposing and creating concrete systems, frameworks, and architectural solutions in digital health security [Purohit et al. 2025, Sami and Toorani 2024, Zala et al. 2024, Kathole et al. 2024, Raghav and Bhola 2023, Chenthara and et al. 2020, Fakeeroodeen and Beeharry 2021, Anusuya and et al. 2022, Watkins and et al. 2023, Tian and et al. 2024, Khaled and Ali 2025].

The remaining contributions are split between two major analytical areas. The studies on Methodology and Conceptual Modeling represent 25% (5 studies), including work focused on defining research approaches, access models, and conceptual design patterns [Churi and Pawar 2024, Kumar et al. 2024, Herwanto and et al. 2024, Monteiro and et al. 2024, Wu et al. 2021]. Finally, studies about Analysis, Evaluation, and Empirical Evidence comprise 20% (4 studies), dedicated to validating existing systems, evaluating algorithms, and gathering empirical data through surveys [Vovk et al. 2021, Tseng and et al. 2025, Madhavi et al. 2024, Dotter and et al. 2025]. Overall, this distribution confirms that the field is primarily focused on engineering novel solutions, with secondary emphasis on formalizing concepts and conducting critical evaluation. In the following subsections, we present the findings that directly address our research questions.

4.2. Clinical Data Anonymization and Reliability Considerations (RQ1)

RQ1 evaluated ”What clinical data anonymization techniques are used in digital health systems?”. The classification of included studies that answers RQ1 is shown in Table 7. The studies were categorized into four intervention categories: Blockchain and Cryptography-Based Solutions, Federated Learning (FL) and Distributed Privacy, Technical and Algorithmic Assessment, and Frameworks and Sociotechnical Aspects.

Thus, it is possible to understand that recent research leans toward technologically disruptive solutions (such as Blockchain and federated learning), but recognizes the need for performance evaluation and the importance of ethical governance for the success of these interventions, constituting the fundamental answer to Research Question RQ1 regarding the clinical data anonymization techniques used in digital health systems. The studies in Table 7 show that anonymization techniques go beyond the removal of direct identifiers, combining classical statistical methods with cryptography, federated learning, and Blockchain. Practical application depends on the type of clinical data, the context of use, and the need to balance privacy, utility, and security.

The included articles emphasize the importance of robust protection in open environments, particularly in cryptography and distributed governance. Studies in category I of Table 7 propose the use of Blockchain technology, Hyperledger Fabric, cryptography with enhanced anonymity models, and pseudonymization in conjunction with access control [Chenthara and et al. 2020, Khaled and Ali 2025]. Furthermore, Blockchain is frequently employed as a security and access control infrastructure, using smart contracts and Zero-Knowledge Proofs (ZK-Snarks) to validate information without revealing the underlying data, acting as a direct anonymization and integrity mechanism [Khaled and Ali 2025].

Several studies address encryption as a central technique for protecting sensitive clinical data in digital healthcare systems [Zala et al. 2024, Raghav and Bhola 2023, Wu et al. 2021]. Blockchain serves as a complementary layer, enhancing the security and

Table 7. Classification of Included Articles by Research Question RQ1.

Intervention Category	**Main Focus**	**Included Articles**
I. **Blockchain and Cryptography-Based Solutions**	Proposals that use blockchain immutability, smart contracts, or cryptography (such as ABE or Homomorphic) to ensure the integrity and access control of anonymized data.	[Sami and Toorani 2024, Zala et al. 2024, Kathole et al. 2024, Raghav and Bhola 2023, Chenthara and et al. 2020, Anusuya and et al. 2022, Wu et al. 2021, Khaled and Ali 2025]
II. **Federated Learning (FL) and Distributed Privacy**	Proposals that use FL or aggregation to enable data analysis (utility) without the raw clinical data leaving the source (distributed anonymization).	[Purohit et al. 2025, Watkins and et al. 2023, Tian and et al. 2024]
III. **Technical and Algorithmic Evaluation**	Studies which focus on the direct application and evaluation of classical anonymization algorithms (e.g., K-Anonymity, generalization, suppression) and their utility trade-offs.	[Vovk et al. 2021, Fakeeroodeen and Beeharry 2021, Tseng and et al. 2025, Monteiro and et al. 2024, Madhavi et al. 2024]
IV. **Frameworks and Sociotechnical Aspects**	Studies that do not propose a technique, but rather frameworks, models, or requirements approaches to ensure privacy, ethics, and trust.	[Herwanto and et al. 2024, Churi and Pawar 2024]

reliability of digital healthcare systems by ensuring that personal and clinical information is stored and transmitted securely, thereby preventing unauthorized access. In the studies analyzed, Blockchain is frequently used as a supporting infrastructure for the anonymization of clinical data [Sami and Toorani 2024, Anusuya and et al. 2022]. It acts as a secure and immutable registry, ensuring the traceability and integrity of information, while also enabling access control and the enforcement of privacy policies via smart contracts. In many cases, Blockchain is combined with classic anonymization techniques, such as generalization, suppression, and encryption, thereby reinforcing data protection without compromising its usefulness [Kathole et al. 2024, Chenthara and et al. 2020].

The Federated learning category (Category II in Table 7) involves training machine learning models on decentralized data without exposing raw clinical records, serving as a distributed anonymization technique [Purohit et al. 2025, Watkins and et al. 2023, Tian and et al. 2024]. Some studies highlight the use of federated learning as an anonymization and privacy-preserving strategy in digital health systems [Purohit et al. 2025, Tian and et al. 2024]. In this model, data remains local to each institution, while only trained parameters or models are shared in aggregate. This architecture allows for collaborative analysis and the construction of predictive models without exposing sensitive individual information, providing a complementary approach to traditional anonymization techniques.

Technical and Algorithmic Evaluation category (Category III in Table 7) details the application of traditional methods and their trade-offs, including suppression, generalization, and randomization (for hybrid algorithms) [Fakeeroodeen and Beeharry 2021]. This category also encompasses the evaluation of concrete anonymization tools and conceptual frameworks (Category IV in Table 7) that guide the application of these techniques through requirements and governance, such as Contextual Integrity [Herwanto and et al. 2024].

4.2.1. Impact of Anonymization Techniques on System Reliability (RQ1.1)

Regarding how these techniques influence the reliability of these systems, RQ1.1 is answered by the same studies that answer RQ1. Reliability is significantly strengthened by employing strategies that ensure structural integrity and robustness. Decentralized architectures, particularly those utilizing Blockchain (e.g., Hyperledger Fabric), eliminate single points of failure, ensuring the integrity and traceability of medical records and protecting against unauthorized alterations [Kathole et al. 2024, Chenthara and et al. 2020]. Data integrity is guaranteed as documents are stored as immutable hash values [Chenthara and et al. 2020].

Furthermore, advanced cryptographic mechanisms, such as Attribute-Based Encryption (ABE) and Zero-Knowledge Proofs (ZK-Snarks), enhance security and reliability by enabling secure data validation and access control without revealing sensitive information [Zala et al. 2024, Anusuya and et al. 2022, Raghav and Bhola 2023]. The Robust and Privacy-Preserving Decentralized Deep Federated Learning (RPDFL) scheme also contributes to reliability by ensuring robustness and resilience against system dropouts during training [Tian and et al. 2024].

Anonymization, particularly when utilizing Blockchain and Federated Learning (FL), substantially enhances system reliability across multiple dimensions, including robustness against failure, guaranteed data integrity, and operational efficiency [Raghav and Bhola 2023]. Strong operational performance metrics, including speed, latency, and throughput, confirm this reliability. Specific examples include blockchain-based systems that demonstrate high storage and transmission efficiency [Kathole et al. 2024, Wu et al. 2021], as well as the Healthchain framework, which achieves low latency (e.g., 2.7 seconds for concurrent updates) when handling large datasets [Chenthara and et al. 2020]. Furthermore, efficient access control models, such as the Risk- and Utility-Based Access Control (RUBAC), contribute to optimizing record retrieval [Churi and Pawar 2024].

Finally, for analytical systems, reliability is measured by maintaining accuracy and data utility post-privacy application. The use of hybrid anonymization algorithms and sanitization methods (e.g., Coati optimization) balances privacy and utility, which are essential for reliable analysis [Fakeeroodeen and Beeharry 2021, Madhavi et al. 2024]. Significantly, the RPDFL scheme demonstrated superior model accuracy and convergence compared to centralized methods (reaching 98% accuracy on some datasets), proving that privacy-preserving techniques can directly enhance the reliability of Machine Learning outcomes [Tian and et al. 2024]. Reliability is also secured by mitigating failure risk in the initial requirements modeling phase through frameworks like Contextual Integrity [Herwanto and et al. 2024].

4.3. Influence of Anonymization on Reliability (RQ2)

Despite the robust technical reliability demonstrated by Blockchain and FL (integrity, low latency), the literature indicates that technical rigor is insufficient for success in healthcare. True dependability requires user engagement and trust. Therefore, RQ2 maps the essential non-algorithmic criteria necessary to preserve patient trust and ensure sociotechnical governance. Some of the included articles reveal that the effectiveness of anonymization in enhancing the reliability of Healthcare Information Systems is assessed through technical rigor, with a focus on maintaining a balance between privacy and utility. This influence occurs by strengthening data integrity and security, which are crucial pillars of reliability. Research question (RQ2) reflects this duality regarding the criteria used to assess the effectiveness of anonymization in preserving user trust. The results of studies that answer RQ2 are shown in Table 8.

Table 8. Evaluation Criteria for Anonymization Effectiveness in Preserving User Trust (RQ2)

Pillar of Effectiveness	Key Criterion / Metric	Focus and Demonstration of Trust
I. Sociopolitical Perception and Public Trust	**Willingness to Share Data**	Measured by trust in public institutions (vs. private companies) and the perceived mitigation of data misuse risk [Dotter and et al. 2025].
	Digital Literacy	A factor that positively influences the public's readiness to engage with and share health data [Dotter and et al. 2025].
	Contextual Integrity (CI)	Prescriptive criterion evaluating moral and political acceptability, ensuring anonymization does not compromise patient autonomy or equity [Kumar et al. 2024].
II. Technical and Algorithmic Robustness	**Privacy Protection Level**	Ability to achieve formal metrics like K-anonymity, L-diversity, and T-closeness, while minimizing information loss [Fakeeroodeen and Beeharry 2021, Monteiro and et al. 2024].
	Privacy-Utility Balance	Ensured by hybrid algorithms, data sanitization, and techniques like Outlier Relocation [Monteiro and et al. 2024, Madhavi et al. 2024].
	Security and Patient-Centric Control	Use of Blockchain for integrity/immutability, ZK-Snarks for validation, and ABAC for total patient control [Chenthara and et al. 2020, Anusuya and et al. 2022, Churi and Pawar 2024].
	Operational Performance /Scalability	Demonstration of practical viability through low latency and high robustness against failures (e.g., dropouts in decentralized FL schemes) [Chenthara and et al. 2020, Tian and et al. 2024].

While some studies focus on objective metrics of techniques such as Re-Identification Risk, Data Distortion, and Utility Preservation for institutional trust, another group prioritizes sociotechnical and governance criteria [Vovk et al. 2021, Kumar et al. 2024]. The latter are essential for user trust and include the formalization of ethical and legal requirements, as well as the mapping of public perceptions of privacy and willingness to share health data [Dotter and et al. 2025, Monteiro and et al. 2024].

The effectiveness of anonymization under the Sociopolitical Perception and Public Trust pillar (Pillar I in Table 8) is principally assessed by the public's willingness to share health data [Dotter and et al. 2025]. The criteria driving this willingness include the level of trust in public institutions (with confidence in German universities and government agencies being higher than in pharmaceutical companies) and the assessment of the severity and probability of data misuse [Dotter and et al. 2025]. Higher confidence in public institutions has a strong and positive influence on the readiness to share data with any organization [Dotter and et al. 2025]. Additionally, digital literacy, measured by experience with video calls, positively affects the disposition to share data with all organizations except pharmaceutical companies [Dotter and et al. 2025].

The Contextual Integrity (CI) framework provides a prescriptive criterion for evaluating effectiveness, focusing on the moral and political acceptability of a practice that may involve anonymization [Kumar et al. 2024]. Evaluation criteria include analyzing how the practice affects individual autonomy or freedom, alters power structures (e.g., exacerbating information asymmetry with insurers), or compromises equality, fairness, and equity [Kumar et al. 2024]. For instance, the automatic transmission of personal fitness information (PFI) to insurers can be perceived as coercive and discriminatory, potentially undermining the ”care” values of the healthcare context [Kumar et al. 2024]. Effectiveness assessment must also consider data accuracy; the automated transmission of unreliable wearable data for medical decision-making may constitute a prima facie violation of contextual integrity [Kumar et al. 2024].

Regarding technical anonymization criteria (Pillar II in Table 8), effectiveness is evaluated by a technique’s ability to achieve a defined level of privacy protection, such as K-anonymity, L-diversity, and T-closeness [Fakeeroodeen and Beeharry 2021, Monteiro and et al. 2024]. A fundamental criterion is the balance between privacy and data utility. Techniques such as Generalization and Hierarchical Generalization are evaluated by their ability to reduce the precision of Quasi-Identifiers (QIs) while protecting identity, while also minimizing information loss [Monteiro and et al. 2024]. Specifically, when handling outliers, the Relocate Outliers technique is effective if the distance between the original and relocated distributions is limited to maintain trustworthiness and prevent an adversary from suspecting intentional modification [Monteiro and et al. 2024]. The effectiveness of data sanitization techniques is also measured by minimizing four side effects: hiding failure, artificial costs, missing costs, and database dissimilarities, which ensures the preservation of quality and privacy [Madhavi et al. 2024].

In Blockchain and Federated Learning (FL) systems, effectiveness is demonstrated through robust security and performance metrics that build user confidence in data security and control. Criteria include ensuring data integrity and immutability (often achieved by storing cryptographic hash values on the chain and encrypted data off-chain, e.g., on IPFS) [Chenthara and et al. 2020]. Confidentiality is evaluated by the use of unique public-key encryption for Electronic Health Records (EHRs) [Chenthara and et al. 2020, Wu et al. 2021]. Crucially, system effectiveness is measured by a patient-centric approach, where the patient retains full control and ownership to grant, read, write, or revoke access permissions using Attribute-Based Access Control (ABAC) and Rules [Chenthara and et al. 2020], [Churi and Pawar 2024].

Advanced anonymity models, such as Zero-Knowledge Proofs (ZK-Snarks), demonstrate effectiveness by allowing medical and insurance agents to obtain information securely and efficiently without revealing the underlying facts [Anusuya and et al. 2022]. Finally, effectiveness in distributed environments is also assessed by scalability (the ability to process large datasets with low latency [Chenthara and et al. 2020]) and robustness against participant dropouts in decentralized FL schemes, which ensures correct protocol execution and gradient security [Tian and et al. 2024].

4.4. Discussion

The SLR confirms that the future of data privacy relies on architectural artifacts, mostly in Blockchain and Federated Learning (FL). These solutions, identified in RQ1, demonstrate superior performance in ensuring data integrity and decentralized control, offering a powerful technical countermeasure to the rising tide of cyberattacks and massive financial losses detailed in the Brazilian context. However, the prevalent focus on *Design and Solution Development* (totaling 55% of the included studies) highlights a critical disconnect. The literature is mature in prototyping robust techniques (RQ1.1) but remains nascent in providing empirical validation of these solutions when deployed at scale within real, heavily regulated national healthcare ecosystems. This gap is the immediate challenge for Information Systems research seeking to inform the Lei Geral de Proteção de Dados (LGPD) implementation.

The most critical finding of this study, stemming from RQ2, is that technical rigor alone is a necessary but insufficient condition for system reliability. By consolidating a dual assessment framework, the study shifts the reliability discourse from measuring technical robustness (e.g., Re-identification Risk) to prioritizing sociotechnical acceptance. The emphasis on Contextual Integrity (CI) proves that an anonymization technique is only effective if its subsequent use aligns with the patient's moral and political expectations. In a high-stakes, regulated domain like healthcare, this means that even an immutable Blockchain record will fail to build trust if the governance model compromises patient autonomy or fairness. This necessity for ethical, relational design highlights the limitations of solutions that treat technology in isolation from its institutional and social contexts.

Given the high stakes in Brazil, the future work must move beyond theoretical proofs of concept to operationalizing the dual criteria. The operational effort requires developing metrics and frameworks capable of objectively quantifying public perception and CI compliance (Pillar I, Table 8) alongside the technical measures (Pillar II). Specifically, longitudinal studies are needed to benchmark the performance of FL and Blockchain using standardized clinical metrics, assessing how they truly impact data utility for AI development without compromising the trust of a digitally variable population. By focusing on this empirical intersection of ethics, technology, and governance, IS research can provide the guidance needed for Brazilian healthcare leaders to design truly dependable digital ecosystems in accordance with the LGPD.

Regarding contributions to the field of Information Systems in Brazil, this study aligns with the GranDSI-BR agenda by providing a consolidated framework for the governance and trustworthy use of sensitive data, which is essential for addressing the Grand Challenge of Ethics and Data Governance in Health. This study addresses an application domain of high social and organizational value: the management of sensitive clinical data in the context of Digital Health**.** This domain, characterized by a growing volume of data and the criticality of information, has a strong intersection with the Grand Research Challenges in Information Systems in Brazil (GranDSI-BR, 2016-2026) [Boscarioli et al. 2017]. This study's contribution aligns with the following challenges: Challenge 2, ”Information Systems and the Challenges of the Open World,” and Challenge 4, ”Sociotechnical Vision of Information Systems” [Boscarioli et al. 2017].

The study aligns with the GranDSI-BR Challenges [Boscarioli et al. 2017] by supporting the secure and ethical use of health data. Specifically, it addresses Challenge 2 (Open World) by examining anonymization as a tool to enhance privacy, security, and reliability, which are attributes crucial for data exchange in interconnected research networks [Boscarioli et al. 2017]. Furthermore, the research tackles Challenge 4 (Sociotechnical Vision) by treating Information Systems as an integration of people and technology. The study emphasizes user trust (RQ2), asserting that a focus on patient confidence, integrity, and legal compliance is necessary to ensure technical anonymization solutions are truly reliable and trustworthy, recognizing that technology is not an isolated artifact [Boscarioli et al. 2017].

4.5. Threats to Validity

This study has some limitations that should be considered when interpreting the findings. First, the search strategy was restricted to four databases, to publications in English and Portuguese, and to the 2020–2025 time window, which may have led to the omission of relevant studies indexed elsewhere, written in other languages, or published in different periods.

Second, the review did not employ backward or forward snowballing, which reduces the chance of capturing additional studies that are less visible to keyword-based searches; future work can complement this protocol with systematic snowballing procedures to broaden the evidence base. Finally, although two reviewers independently applied the inclusion, exclusion, and quality criteria and resolved disagreements by consensus, some degree of subjectivity in study selection, classification, and interpretation of results is unavoidable.

5. Conclusion

This Systematic Literature Review confirms that anonymization significantly enhances the reliability of Healthcare Information Systems, an objective met by mapping dominant techniques and their influence on trustworthiness. The analysis demonstrates that interventions such as Blockchain and Federated Learning (FL) positively enhance data integrity and ensure the critical balance between privacy and utility. Furthermore, the study established a dual assessment framework where effectiveness is not solely measured by technical metrics (e.g., Re-identification Risk) but also by sociotechnical criteria such as Contextual Integrity and the public's perception of trust in institutional governance. This consolidation addresses the primary challenge identified: moving the field beyond purely algorithmic solutions to holistic, dependable systems, meeting the GranDSI-BR Challenge of Ethics and Data Governance in Health.

Despite these contributions, this study is subject to limitations inherent in the SLR methodology, including the temporal scope (2020–2025) and the constraint of the defined search string. Future work could prioritize three avenues. First, benchmarking current FL and Blockchain frameworks using standardized metrics to compare real-world performance against claimed reliability, and second, developing a formal sociotechnical framework that operationalizes Contextual Integrity and public trust criteria for system design. Finally, expanding the SLR's scope to include terms related to user acceptance models and specific organizational case studies would further validate the alignment between technical effectiveness and successful clinical adoption.

According to the Systematic Literature Review conducted in this study, it is concluded that anonymization improves the reliability of Healthcare Information Systems. Intervention techniques, such as Blockchain and Federated Learning (RQ1), positively influence this reliability by strengthening integrity and ensuring a balance between privacy and data usefulness for secondary purposes (RQ1.1). However, the effectiveness of this anonymization is assessed from two perspectives (RQ2). In addition to technical metrics (such as Re-Identification Risk), sociotechnical criteria are also evaluated, including ethical governance, Contextual Integrity, and the patient's perception of trust in the system and how data is stored. This dual evaluation demonstrates that this perception depends on the ability to sustain relational trust in the system.

Acknowledgments

We thank Google Gemini for assistance in language revision.

References

Anusuya, R. and et al. (2022). Privacy-preserving blockchain-based ehr using zk-snarks. In Raman, I. and et al., editors, *Computational intelligence, cyber security and computational models*. Springer.

Boscarioli, C., Araujo, R. M., and Maciel, R. S. P. (2017). *I GranDSI-BR – Grand Research Challenges in Information Systems in Brazil 2016–2026*. Special Committee on Information Systems (CE-SI), Brazilian Computer Society (SBC).

Camêlo, M. and Alves, C. (2023). G-priv: Um guia para apoiar a especificação de requisitos de privacidade em conformidade com a lgpd. *iSys - Brazilian Journal of Information Systems*, 16.

Carvalho, L. P., Oliveira, J., Santoro, F. M., and Cappelli, C. (2021). Social network analysis, ethics and lgpd, considerations in research. *iSys - Brazilian Journal of Information Systems*, 14(2):28–52.

Chenthara, S. and et al. (2020). Healthchain: a novel framework on privacy preservation of electronic health records using blockchain technology. *PLOS ONE*.

Churi, P. and Pawar, A. (2024). Rubac: proposed access control for flexible utility–privacy model in healthcare. *SN Computer Science*, 5:297.

Dotter, C. and et al. (2025). Sharing health data for research purposes: results of a population survey in germany. *BMC Health Services Research*, 25:699.

Fakeeroodeen, Y. N. and Beeharry, Y. (2021). Hybrid data privacy and anonymization algorithms for smart health applications. *SN Computer Science*, 2(126).

Herwanto, G. B. and et al. (2024). Integrating contextual integrity in privacy requirements engineering: a study case in personal e-health applications. In Phillipson, F., Eichler, G., Erfurth, C., and Fahrnberger, G., editors, *Innovations for Community Services. I4CS 2024*, Communications in Computer and Information Science, v. 2109. Springer, Cham.

Kaspersky (2025). Novo estudo mostra aumento das vítimas de ransomware no Brasil. https://www.kaspersky.com.br/about/press-releases/kaspersky-novo-estudo-mostra-aumento-das-vitimas-de-ransomware-no-brasil.

Kathole, A. B., Patil, S. D., Kumbhare, S., and et al. (2024). Electronic health records protection strategy by using blockchain approach. *Multimedia Tools and Applications*, 83:86883–86894.

Khaled, O. and Ali, A. F. (2025). Using blockchain and smart contracts to secure medical data management system. In Abdelgawad, A., Jamil, A., and Hameed, A. A., editors, *Intelligent systems, blockchain, and communication technologies. ISBCom 2024*, Lecture Notes in Networks and Systems, v. 1268. Springer, Cham.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical report, Keele University and University of Durham.

Kumar, P. C., Zimmer, M., and Vitak, J. (2024). A roadmap for applying the contextual integrity framework in qualitative privacy research. In *Proceedings of the ACM on Human-Computer Interaction*, volume 8, pages 1–29.

Madhavi, M., Sasirooba, T., and Kumar, G. K. (2024). Securing sensitive medical information with basic and pre-large coati optimization algorithm for e-health system data sanitation. *Wireless Personal Communications*, 136:1261–1281.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Annals of Internal Medicine*, pages 264–269.

Monteiro, M. and et al. (2024). Patterns of data anonymization. In *Proceedings of the 29th European Conference on Pattern Languages of Programs, People, and Practices (EuroPLoP '24)*, pages 1–9, New York. Association for Computing Machinery.

Mumford, E. (2006). The story of socio-technical design: Reflections on its successes, failures and potential. *Inf. Syst. J.*, 16:317–342.

Ponemon Institute and IBM (2025). Relatório do custo das violações de dados 2025: A lacuna na supervisão da ia (resumo executivo). https://brasil.newsroom.ibm.com/2025-07-30-Relatorio-da-IBM-Custo-medio-de-uma-violacao-de-dados-no-Brasil-atinge-R-7,19-milhoes.

Purohit, R. M., Verma, J. P., Jain, R., and et al. (2025). Fedblocks: federated learning and blockchain-based privacy-preserved pioneering framework for iot healthcare using ipfs in web 3.0 era. *Cluster Computing*, 28:139.

Queiroz, M. J., Lino, N. C. Q., and Motta, G. H. M. B. (2016). Uma ontologia de domínio para preservação de privacidade em dados publicados pelo governo brasileiro. In *Anais do XII Simpósio Brasileiro de Sistemas de Informação (SBSI)*, pages 009–016, Florianópolis, SC. Sociedade Brasileira de Computação (SBC).

Raghav, N. and Bhola, A. K. (2023). Healthcare framework for privacy-preserving based on hyperledger fabric. In Marriwala, N., Tripathi, C., Jain, S., and Kumar, D., editors, *Mobile Radio Communications and 5G Networks*, Lecture Notes in Networks and Systems, v. 588. Springer, Singapore.

Sami, K. T. and Toorani, M. (2024). Blockchain-based access control for electronic health records. In Abie, H., Gkioulos, V., Katsikas, S., and Pirbhulal, S., editors, *Secure and*

Resilient Digital Transformation of Healthcare. SUNRISE 2023, Communications in Computer and Information Science, v. 1884. Springer, Cham.

Sposito, S. L., Sales, R. d. S., Canedo, E. D., and Silva, G. R. S. (2024). An anonymization library for rapid and diverse anonymization of brazilian personal data. In *Concurso De Trabalhos De Conclusão De Curso Em Sistemas De Informação - Simpósio Brasileiro De Sistemas De Informação (SBSI '24)*, pages 192–201, Porto Alegre. Sociedade Brasileira de Computação.

Tian, Y. and et al. (2024). Robust and privacy-preserving decentralized deep federated learning training: focusing on digital healthcare applications. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 21(4):890–901.

Tseng, F. P. and et al. (2025). Patient privacy information retrieval with longformer and crf, followed by rule-based time information normalization: a dual-approach study. In Jonnagaddala, J., Dai, H. J., and Chen, C. T., editors, *Large Language Models for Automatic Deidentification of Electronic Health Record Notes. IW-DMRN 2024*, Communications in Computer and Information Science, v. 2148. Springer, Singapore.

Vovk, O., Piho, G., and Ross, P. (2021). Evaluation of anonymization tools for health data. In Bellatreche, L., Chernishev, G., Corral, A., Ouchani, S., and Vain, J., editors, *Advances in Model and Data Engineering in the Digitalization Era. MEDI 2021*, Communications in Computer and Information Science, v. 1481. Springer, Cham.

Watkins, M. and et al. (2023). Privacy-preserving data aggregation scheme for e-health. In Al-Sharafi, M. A. and et al., editors, *Proceedings of the 2nd International Conference on Emerging Technologies and Intelligent Systems. ICETIS 2022*, Lecture Notes in Networks and Systems, v. 573. Springer, Cham.

Wohlin, C., Runeson, P., Host, M., Ohlsson, M. C., Regnell, B., and Wesslen, A. (2012). *Experimentation in Software Engineering*. Springer, Berlin, Heidelberg.

Wu, H., Dwivedi, A. D., and Srivastava, G. (2021). Security and privacy of patient information in medical systems based on blockchain technology. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 17(2s):1–17.

Zala, K., Thakkar, H. K., Dholakia, N., and et al. (2024). Designing an attribute-based encryption scheme with an enhanced anonymity model for privacy protection in e-health. *SN Computer Science*, 5:203.