technologies

MDPI

Review

Heuristic Techniques for Assessing Internet Privacy: A Comprehensive Review and Analysis

David Cevallos-Salas *, José Estrada-Jiménez and Danny S. Guamán

Departamento de Electrónica, Telecomunicaciones y Redes de Información, Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, Quito 170525, Ecuador; jose.estrada@epn.edu.ec (J.E.-J.); danny.guaman@epn.edu.ec (D.S.G.)

* Correspondence: david.cevallos03@epn.edu.ec

Abstract

While Internet privacy is a subjective term that is challenging to define, describe, and quantify, assessing the level of privacy provided by data processors offering services over the Internet is essential for detecting privacy flaws and enabling continuous improvement. Moreover, assessing Internet privacy is fundamental for estimating the risk of personal data disclosure, the degree of compliance with privacy regulations, and the effectiveness of implemented protection mechanisms. Remarkably, the absence of a standardized criterion for this assessment has led to the proliferation of diverse heuristic techniques applied with different approaches. In this paper, we conduct an in-depth analysis and introduce a novel taxonomy for categorizing existing heuristic techniques to assess Internet privacy. Moreover, we scrutinize various protection mechanisms designed to enhance users' privacy. We cover this broad topic across all domains of application and levels of automation, considering all relevant papers regardless of publication year, ultimately providing a comprehensive review of this important field of knowledge. Leveraging our proposed classification framework, we systematically organize and categorize 160 papers carefully selected from 934 candidates, elucidating existing gaps and challenges while foreseeing future research directions. Overall, our findings reveal that most studies predominantly rely on information measurement methods for assessing Internet privacy. Although most heuristic techniques are based on automatic mechanisms, they are applied with a clear focus on the traditional use of Internet services through a web browser, demanding more research efforts for other domains. The development of new technologies that incorporate privacy-by-default and include telemetry modules in their architectures will be essential for assessing and enhancing users' privacy when delivering services over the future Internet.

Keywords: Internet privacy; privacy assessment; heuristics; metric; protection mechanism; review

Academic Editor: Arash Habibi Lashkari

Received: 14 July 2025
Revised: 26 July 2025
Accepted: 30 July 2025
Published: 22 August 2025

Citation: Cevallos-Salas, D.; Estrada-Jiménez, J.; Guamán, D.S. Heuristic Techniques for Assessing Internet Privacy: A Comprehensive Review and Analysis. *Technologies* **2025**, *13*, 377. https://doi.org/10.3390/technologies13090377

1. Introduction

The Internet has revolutionized the delivery of services on a global scale [1]. Initially developed as the ARPANET network in the late 1960s, its primary goal was to facilitate communication among universities and research institutions within the United States, presenting technical limitations for global scalability [2]. Moreover, mechanisms for privacy protection and secure communications were not prioritized in its original architecture, protocol stack, and operational framework [3].

As the Internet started to gain widespread adoption, addressing these scalability shortcomings became increasingly critical [4]. However, several foundational technologies that

enabled subsequent advancements were similarly developed without integrating privacy and security considerations [2]. For instance, a key development was the introduction of cookies in 1994, driven by the stateless request-response model of the Hypertext Transfer Protocol (HTTP), which remains widely used in Internet communications today [5]. Cookies are files stored by websites in users' web browsers and contain session-related data, such as authentication credentials, users' preferences, and other state information [4]. While this technology enabled stateful interactions and enhanced users' experience, it also introduced new privacy concerns by facilitating personal data exfiltration and enabling persistent tracking of users, thereby worsening the overall Internet privacy landscape [2]. Therefore, in response to the wide range of threats to Internet privacy, various protection mechanisms have been proposed during the last decade, ranging from detection to advanced response and recovery solutions [6].

However, protecting privacy in modern Internet communications presents major challenges. The Internet has evolved far beyond the traditional paradigm of users accessing content solely through web browsers [7]. Today, substantial volumes of information are exchanged via web services and Application Programming Interfaces (APIs) without the involvement of a graphical interface [8]. Mobile applications running on smartphones communicate vast amounts of users' personal data to remote servers for delivering services [9]. The advent of the Internet of Things (IoT) has enabled the interconnection of everyday devices with services that were once only accessible via web browsers [10]. Similarly, Industrial Control Systems access and communicate personal data through the Internet for operational purposes [11]. These developments have significantly amplified the exposure of users to potential risks such as identity theft, harassment, and cyber fraud [12]. The main question is: What level of Internet privacy does a data processor provide when delivering a service to users in order to mitigate these risks?

The term Internet privacy is inherently subjective and difficult to define, with no universally accepted framework or consensus on how to describe or quantify it [13]. Consequently, researchers have proposed various techniques to support Internet privacy assessment [14], each of which relies on a specific metric to conduct the evaluation. While some metrics provide only qualitative insights, others enable quantification and the definition of ordering criteria, ultimately supporting different levels of assessment [15]. However, since Internet privacy cannot be directly assessed, the metric is subsequently analyzed, interpreted, and logically linked to a corresponding level of Internet privacy [4]. In the context of this research, these techniques are therefore considered heuristic, even if grounded in strong mathematical foundations, as they estimate a level of Internet privacy. Furthermore, all existing techniques remain inherently heuristic, as none can fully capture the complexity of Internet privacy or the broad spectrum of associated threats, thereby comprising simplified solutions to the problem [16].

Assessing Internet privacy is of paramount importance, as it enables the estimation of the risk of personal data disclosure, the degree of compliance with privacy regulations, and the effectiveness of implemented protection mechanisms [17]. However, the significant challenge of precisely defining what to assess and to what extent has driven the development of diverse solutions featuring various approaches and levels of automation [12]. Therefore, several heuristic techniques have been employed for Internet privacy assessment, spanning from manual methods such as surveys and questionnaires to fully automatic approaches leveraging algorithms and machine learning solutions [1].

Regardless of the wide range of proposed solutions and approaches, research efforts generally aim to assess either the existing level of Internet privacy within an information system or the change resulting from the implementation of a protection mechanism, typically by comparing the initial and final states of the system [18]. On the other hand, certain

protection mechanisms, such as privacy malware classifiers, Intrusion Detection Systems (IDS), Data Loss Prevention (DLP) solutions, and Intrusion Prevention Systems (IPS), allow for a direct evaluation of their effectiveness [1]. In such cases, the achieved effectiveness is interpreted as the final level of Internet privacy. This particular approach is referred to in this research as the Protection Mechanism Approach (PMA).

1.1. Problem Statement

Given that Internet privacy is a subjective term that is difficult to define, describe, and quantify, it has prompted researchers to develop a variety of heuristic techniques for its assessment. These techniques, lacking a comprehensive taxonomy for classification, are applied through solutions with varying levels of assessment and automation across diverse domains. Similarly, numerous protection mechanisms designed to enhance Internet privacy have been proposed without a systematic organization. Consequently, there is an urgent need to systematically identify, classify, and analyze this knowledge to reveal existing gaps and challenges, thereby guiding future research directions.

This would provide a clear understanding of the available heuristic techniques for assessing the level of privacy offered by data processors when delivering services over the Internet, along with the protection mechanisms in place to detect, prevent, contain, and respond to personal data leaks. Numerous cases of personal data exfiltration worldwide have demonstrated why such understanding is essential by offering important lessons learned. For example, Episource, a healthcare services company, suffered a data breach in February 2025 that affected more than 5.4 million people [19]. Attackers accessed sensitive personal and medical information, including social security numbers, medical records, and insurance details [20]. Similarly, in 2024, the massive RockYou2024 database containing approximately 9.9 billion unique plaintext passwords collected from various websites and online platforms through brute-force attacks was released [21]. In 2023, the cybercriminal group CL0P exploited the zero-day vulnerability CVE-2023-34362 in the MOVEit Transfer software, gaining unauthorized access to sensitive data [22]. Additionally, the Conti ransomware attack on the Costa Rican government in 2022 triggered a national emergency, exposing hundreds of personal records belonging to Costa Rican citizens who had entrusted their data to public institutions [23]. All these case studies resulted in widespread personal data leakage affecting numerous organizations and individuals within a short time frame, and could likely have been prevented, or their impact mitigated, if the low level of Internet privacy of the affected services had been assessed and appropriate protection mechanisms implemented.

1.2. Contributions

The major contributions of this review are as follows:

- We propose a novel taxonomy to identify and classify existing heuristic techniques used for assessing Internet privacy.
- We suggest a comprehensive classification scheme that takes into consideration the categories of heuristic techniques, levels of assessment, levels of automation, types of protection mechanisms, and broad domains. This framework aims to systematize and organize existing knowledge about the assessment of Internet privacy and proposed protection mechanisms.
- We offer a comprehensive technical background that summarizes and explains the key metrics employed by existing heuristic techniques to assess Internet privacy.
- We identify gaps and challenges that require attention.
- We present potential future research directions for assessing Internet privacy and developing effective protection mechanisms.

1.3. Paper Organization

The rest of the paper is structured as follows: Section 2 provides an overview of previous related work, while Section 3 offers a complete technical background to ensure comprehension of essential terminology and concepts. Section 4 outlines the methodology used in this review. Section 5 is dedicated to presenting and analyzing the results achieved, and Section 6 presents a comprehensive discussion addressing the identified gaps, challenges, and future research directions. Section 7 focuses on the analysis of potential threats to validity that have been carefully considered, demonstrating the validity of our results. Section 8 exposes the conclusion of the review.

2. Related Work

Numerous researchers have made substantial contributions through comprehensive literature reviews on Internet privacy. The majority of these secondary studies focus on summarizing and analyzing solutions that describe potential threats and propose protection mechanisms within specific domains. However, only a limited number of researchers have focused on reviewing solutions that address the challenge of assessing Internet privacy through heuristic techniques and their associated metrics. This section highlights some of this prior important work, covering both scenarios.

Reviews related to Internet Privacy Concerns (IPCs) as an assessment of Internet privacy: Surveys are one of the primary tools used by researchers to assess the level of Internet privacy. In the comprehensive review conducted by Yun et al. [24], common strategies for gathering data on IPCs through surveys are analyzed and linked to a level of Internet privacy. Furthermore, Bartol et al. [25] present a review of how these techniques are applied for assessing Internet privacy, along with the challenges associated with survey-based studies. In a related study, the same authors [26] review and highlight the importance of selecting appropriate scales when conducting surveys in the context of social networks and how these scales can be interpreted as levels of Internet privacy. Collectively, these studies aim to systematize existing knowledge on the analysis of IPCs to enable the assessment of Internet privacy through a logical interpretation of the obtained results. Table 1 explains the common biases associated with survey-based assessment techniques that must be considered when conducting a research using this approach.

Analyzing privacy policies: Due to the enactment of privacy regulations, such as the European General Data Protection Regulation (GDPR), privacy policies have gained an important role in describing how users' personal data are collected, processed, and treated. Schyff et al. [27] present a review of the techniques used for analyzing these legal documents and propose a research agenda to improve them. Del Álamo et al. [28] offer a similar approach in their review, highlighting that only 15% of existing contributions until early 2022 propose a privacy metric when analyzing privacy policies.

Anonymization widely deployed: Anonymization is a fundamental protection mechanism for Internet communications. The review conducted by Majeed et al. [29] illustrates how this concept is applied in graph-based clustering mechanisms, summarizing several existing solutions. It also provides a list of metrics, such as entropy leakage and statistical indicators, that are valuable for interpreting and assessing the level of Internet privacy achieved by an anonymized information system.

Internet advertising and Online Social Networks (OSN): An interesting approach reviewing the existing literature on methods to assess the level of Internet privacy about advertising is presented by Barajas et al. [30]. In this tutorial, the success rate of tests conducted over several studies in a period of up to two months is used as a criterion of Internet privacy.

Table 1. Common biases associated with survey-based assessment techniques.

Bias	Explanation
Social desirability	This occurs when participants respond in a way they believe will be viewed favorably by others, rather than providing honest answers [26].
Sample representativeness	Refers to how well the participants in a study reflect the characteristics of the broader population being studied. If the sample is not representative (e.g., skewed by age, gender, or tech use), the findings about Internet privacy perceptions or behaviors may not generalize to the full population [25].
Questions' order	This bias occurs when the sequence of questions influences how participants respond. Earlier questions can frame or prime thoughts that affect answers to later ones, potentially distorting the true opinions or behaviors related to Internet privacy [31].
Questions' formulation	Happens when the wording or phrasing of a question influences how participants interpret and answer it. Leading, vague, or emotionally charged language can skew responses, affecting the reliability of Internet privacy findings [31].
Context	Arises when the surrounding environment, situation, or framing of the study influences participants' responses. Participants' mood is also within this category [32].
Memory	Occurs when participants inaccurately recall past events or behaviors. In privacy studies, this can lead to unreliable data if people forget, distort, or misremember their actions related to data sharing or Internet privacy settings [26].
Central tendency	This bias occurs when participants tend to avoid extreme responses and instead choose middle or neutral options [32].
Nonresponse	Occurs when people skip some questions or completely refuse to participate in the study, and their absence causes the results to not fully represent the entire population, especially if those who do not respond have different Internet privacy attitudes [26].

The survey conducted by Esposito et al. [33] evaluates some solutions for detecting malicious activities and dishonest users in OSN. The metrics summarized from the analyzed papers serve as indicators of the reliability and levels of Internet privacy achieved. Similarly, with a strong focus on social networks, the state-of-the-art study proposed by Eke et al. [34] consolidates various privacy solutions that revolve around protection mechanisms designed to restrict the inclusion of personal data in user profiles, employing machine learning classification metrics for their evaluation, and linking them to levels of Internet privacy under the PMA approach. A similar work is offered by the review of Senette et al. [35], which analyzed modern techniques to avoid user identity linkage in OSN.

Mobile broad domain: Mobile applications have gained widespread use for the delivery of services over the Internet. Desmal et al. [36] present a detailed review of mobile application satisfaction models, with a particular focus on Internet privacy in the context of Mobile Government (mGovernment) portals. Additionally, the authors introduce a comprehensive 6-item evaluation model, where users' perceptions of privacy and security play a central role in this framework to determine a level of Internet privacy.

Notably, the IoT has made it possible to deliver services through interconnected devices and sensors that can collect and transmit personal data for various applications and industries. A general overview of Internet privacy within this domain is provided in [37]. Additionally, a comprehensive analysis of protection mechanisms for safeguarding Internet privacy in the context of IoT is offered by Haque et al. [38]. While these reviews are valuable by providing an extensive examination of security contributions in IoT and emphasizing the importance of Internet privacy, they do not explore specific techniques for assessing Internet privacy in this domain. On the contrary, Dong et al. [39] contribute with a detailed explanation of metrics for assessing Internet privacy, with a particular focus on IoT, proposing a novel metric for assessing the utility-privacy relationship. The review conducted by Bhattacharjya et al. [40] offers some techniques to provide privacy protection

in the domain of Internet of Underwater Things (IoUT) through protection mechanisms based on federated learning.

Finally, due to IDS protection play a significant role, the review presented in [41] makes a comparative analysis of various machine learning algorithms conducted to assess their capabilities for intrusion detection in IoT systems, suggesting how their performance is interpreted and linked to a level of Internet privacy under the PMA approach.

Healthcare and education sectors: There are also noteworthy reviews dedicated to summarizing solutions aimed at enhancing Internet privacy within the healthcare sector. For instance, Al-Qarni [42] provides a concise overview of the primary threats faced by this sector, particularly those stemming from privacy malware. The author underscores the need for implementing robust protection mechanisms, although no specific metric is proposed for assessing their effectiveness. On the other hand, Hameed et al. [43] present a literature review of contributions related to security and privacy in the healthcare sector within the context of the IoT, mainly analyzing the effectiveness of solutions based on machine learning algorithms and defining levels of Internet privacy using the PMA approach.

Moreover, the Internet has not only revolutionized access to information but has also given rise to online education, where vulnerable users such as children and teenagers share significant amounts of personal data. In this context, Kumar and Knox [44] have conducted a comprehensive review that consolidates and examines solutions related to taxonomies and protection mechanisms for online education. Nevertheless, the review lacks a specific metric for evaluating the effectiveness of these protection mechanisms or determining the level of Internet privacy they can attain.

Privacy in storage systems for delivering services over the Internet: Preserving Internet privacy through storage protection mechanisms is of critical importance for service delivery. In their methodological survey, Rajkumar and Dhanakoti [45] conduct a comparative analysis of existing techniques for safeguarding information stored in the cloud, outlining their respective advantages and limitations. The research demonstrates how these mechanisms can enhance Internet privacy in APIs, shared servers, and backup services. However, the study does not propose specific metrics for interpreting or quantifying the level of Internet privacy achieved through these protection mechanisms.

Table 2 provides a comparison among the previous related work and our review. It details whether each work (a) tackles or not papers that analyze the operation of protection mechanisms; and (b) exposes papers that analyze heuristic techniques for assessing Internet privacy using the PMA and/or any other approach. The domain of application of each study is also indicated in the table. To the best of our knowledge, this is the first study that conducts a review of heuristic techniques for assessing Internet privacy and protection mechanisms completely domain agnostic.

Table 2. A comparison of previous related work with this review.

Reference	Year	Protection Mechanisms	Heuristic Techniques		Domain
			PMA	Others	
Bhattacharjya et al. [40]	2025	✓		✓	IoUT
Schyff et al. [27]	2024			✓	Web
Senette et al. [35]	2024	✓	✓		Social networks
Al-Qarni [42]	2023	✓			Healthcare
Bartol et al. [26]	2023			✓	Social networks
Benamor et al. [41]	2023	✓	✓		Web

Table 2. *Cont.*

Reference	Year	Protection Mechanisms	Heuristic Techniques		Domain
			PMA	Others	
Esposito et al. [33]	2023	✓		✓	Social networks
Kumar and Knox [44]	2023	✓			Online education
Rajkumar and Dhanakoti. [45]	2023	✓			Web
Barajas et al. [30]	2022	✓	✓		Web
Del Álamo et al. [28]	2022			✓	Web and Mobile
Desmal et al. [36]	2022	✓	✓		Mobile
Haque et al. [38]	2022	✓			IoT
Majeed et al. [29]	2022	✓		✓	Web
Bartol et al. [25]	2021			✓	Web
Hameed et al. [43]	2021	✓	✓		Healthcare IoT
Eke et al. [34]	2019	✓	✓		Social networks
Dong et al. [39]	2018	✓		✓	IoT
Youm [37]	2017	✓			IoT
Yun et al. [24]	2014			✓	Web
This research	2025	✓	✓	✓	All

3. Technical Background

In this section, we provide a comprehensive explanation of the key concepts to support a thorough understanding of the review. Specifically, we detail the proposed taxonomy of heuristic techniques for assessing Internet privacy. In addition, we analyze related concepts including levels of assessment, levels of automation, protection mechanisms, and broad domains.

3.1. Heuristic Techniques

Figure 1 introduces our proposed taxonomy for categorizing the heuristic techniques used for assessing Internet privacy. A technique encompasses any method, procedure, or approach that uses a metric to enable interpretation, analysis, and association with a level of Internet privacy. Our taxonomy comprises six categories in total.

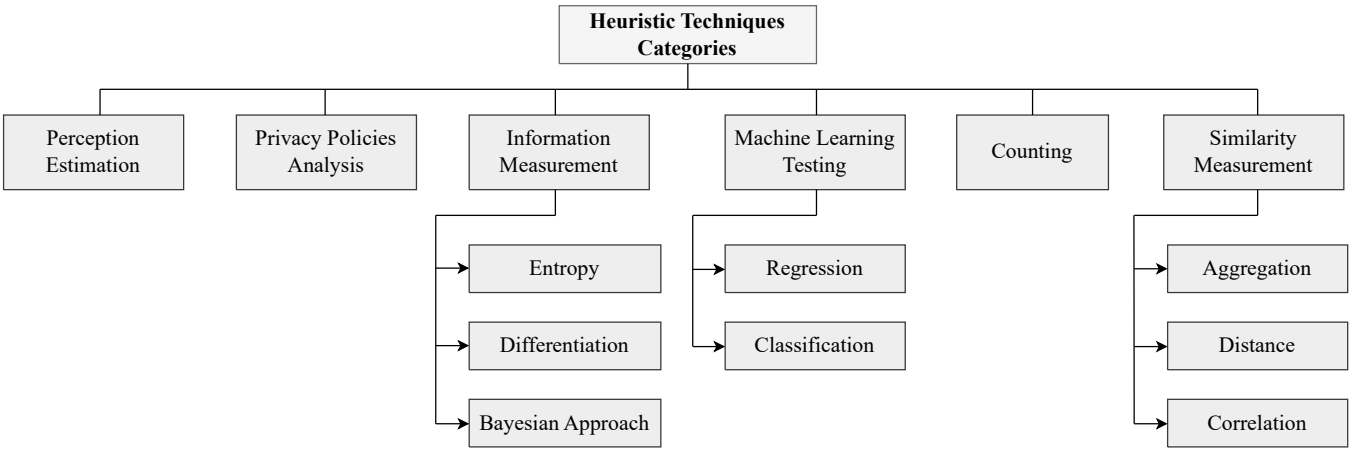

Figure 1. Proposed taxonomy of heuristic techniques for assessing Internet privacy.

3.1.1. Perception Estimation

Perception refers to the cognitive ability of individuals to interpret sensory information received through their senses. In the context of Internet services, users form personal assessments of privacy based on their experiences and interactions [46]. Techniques in this category are particularly valuable for estimating the risk of personal data disclosure by evaluating the amount of sensitive data users handle and the potential impact of a data breach. Common approaches include yes/no surveys to assess whether users feel a sense of privacy when using Internet services, questionnaires to identify the categories of personal data most frequently handled by users, and Likert scale-based evaluations to gauge the extent to which users believe data processors protect their personal data. A key advantage of perception-based techniques is their flexibility in data collection, ranging from traditional surveys to advanced methods such as emotion recognition through human-computer interaction for further analysis. However, as with other cognitive processes, perception-based assessments are highly susceptible to bias and sensory limitations, which may lead to misleading conclusions [31]. Additionally, these techniques primarily reflect subjective opinions [47].

3.1.2. Privacy Policies Analysis

Privacy policies establish guidelines for the collection, use, and handling of personal data by data processors [48]. Therefore, they serve as a valuable proxy for assessing the level of Internet privacy an entity may offer, based on the analysis of the information disclosed in these legal documents. This analysis is mainly useful for evaluating the degree of compliance with privacy regulations. Common techniques in this category include identifying privacy practices and mapping them to specific legal principles, assessing the readability and comprehensibility of the document, and measuring the length and scope of the privacy policy [49].

One significant advantage of these techniques is their flexibility in collecting privacy policies information, ranging from manual compilation to automatic web crawlers [50]. Moreover, these techniques are objective and may lead to more reliable conclusions with a lower likelihood of errors compared to perception estimation. However, the techniques within this category have some disadvantages. Firstly, many Internet services do not regularly update their privacy policies, and some may not have one at all [11]. Assessing how well the information in the privacy policy aligns with reality is another critical consideration [51]. Additionally, the diverse data formats in which privacy policies are presented over the Internet can complicate the automatic collection of information from them [52].

3.1.3. Information Measurement

The techniques within this category are among the most important for assessing Internet privacy. Many authors consider the amount of information disclosed by a system as a strong indicator of its level of Internet privacy [53]. However, it is important to emphasize that information measurement alone does not constitute a level of Internet privacy without a subsequent interpretation and analysis. Particularly, these techniques are useful for evaluating the effectiveness of a protection mechanism. Within this category, we identify three key subcategories for measuring information: Entropy-based techniques, Differentiation-based techniques, and Bayesian approaches.

Entropy: In information theory, entropy is used as a measure of uncertainty, and it comprises the most fundamental concept for measuring information [54]. Consider the scenario in Figure 2 where a data processor handles up to four attributes while offering a service over the Internet for three users through its information system. An information system is a set of elements that collect, use, treat, or transmit users' personal data over

the Internet. An attribute comprises any category of personal data that could serve as an identifier (e.g., full name, passport number, among others), quasi-identifier (e.g., date of birth, job title, gender, among others), or sensitive information (e.g., financial or academic records). If access is gained to these original attributes, an attacker could potentially identify the users and steal their personal data. In contrast, a privacy-aware information system would protect each attribute (by applying a protection mechanism by privacy-by-default, such as masking, noise addition, among others), ultimately creating multiple protected versions for each attribute. If an attacker gathers one of these original attributes (now protected), the likelihood of identifying the user is reduced due to increased uncertainty. Therefore, a higher entropy implies less information disclosure, and thus a higher level of Internet privacy. Conversely, a lower entropy indicates that the information system reveals more information, involving a lower level of Internet privacy. The most representative technique for quantifying this uncertainty is to calculate the Shannon entropy, typically expressed in bits, which consists of computing the probability of gathering each protected attribute of the information system based on its frequency of appearance.

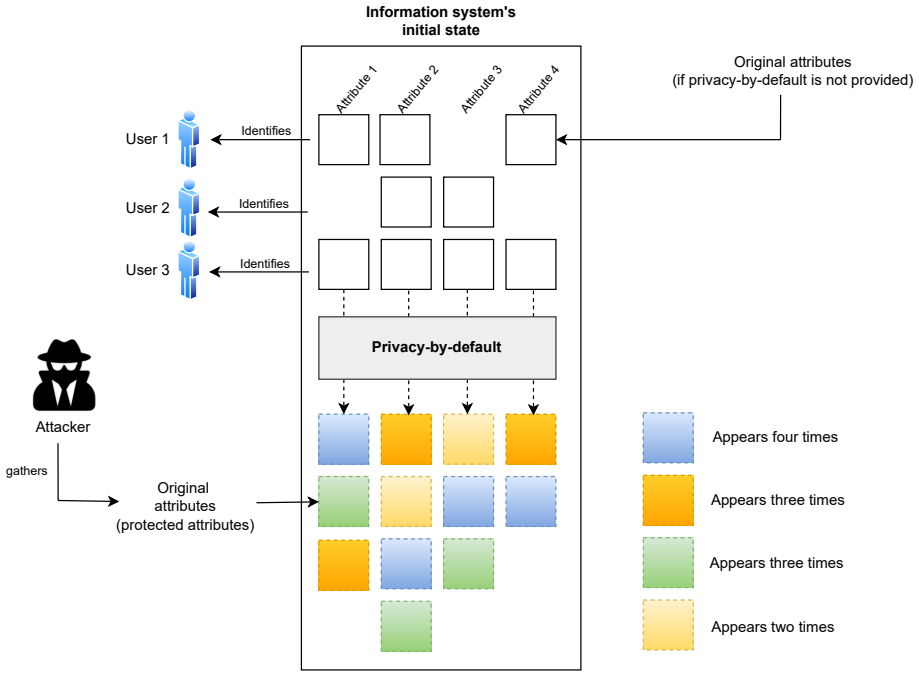

Figure 2. Information system with a privacy-by-default approach.

Thus, for example, the information system of Figure 2 can represent any of the four attributes from three users through a total of four protected attributes, each one represented by a different color that can appear several times. If this information system is defined by the random variable X, with K denoting the total number of different protected attributes ($K = 4$), each one with a probability of appearance equal to $P(X = k)$, where k is one of the possible K protected attributes, then the Shannon entropy of X, denoted as $H(X)$, is defined by Equation (1) as follows [55,56]:

$$H(X) = -\sum_{k=1}^{K} P(X = k)\, log_2[P(X = k)] \quad (1)$$

The probabilities of each k protected attribute can therefore be calculated based on its frequency of appearance. For instance, the first protected attribute represented in blue appears four times out of the twelve possible cases, achieving a probability value of $P(X = 1) = \frac{1}{3}$. The second and third protected attributes, represented in orange and blue,

respectively, each appear three times out of twelve cases, achieving the same probability value of $P(X = 2) = P(X = 3) = \frac{1}{4}$. The last protected attribute represented in yellow just appears two times, reaching a probability value of $P(X = 4) = \frac{1}{6}$. In this way, Equation (1) can be solved as Equation (2) details:

$$H(X) = -\frac{1}{3}\, log_2\left(\frac{1}{3}\right) - \frac{1}{4}\, log_2\left(\frac{1}{4}\right) - \frac{1}{4}\, log_2\left(\frac{1}{4}\right) - \frac{1}{6}\, log_2\left(\frac{1}{6}\right) \approx 1.96\ [bits] \quad (2)$$

The maximum Shannon entropy of this information system is equal to 2 bits and is achieved when all protected attributes are equally probable. Therefore, the achieved Shannon entropy value of 1.96 bits is high, and the information system provides a high level of Internet privacy. Wu et al. [55] and Motahari et al. [56] offer more examples of how to calculate this metric and interpret it in terms of Internet privacy.

Differentiation: Of particular importance is assessing the level of Internet privacy achieved after the implementation of an external protection mechanism that adds more uncertainty to the original attributes protected by privacy-by-default or, if the information system is not privacy-aware, provides a first layer of protection to these attributes. For this purpose, various differentiation techniques can be applied that consider the information system's states both before (initial state) and after (final state) the implementation of the protection mechanism, as Figure 3 depicts.

A prominent technique within this category is to measure the Shannon Information Gain. Shannon Information Gain is a metric that calculates the difference between the initial and final Shannon entropy of an information system. A negative value of Shannon Information Gain indicates a loss of information, increasing Shannon entropy and, consequently, achieving a higher level of Internet privacy [57]. This aligns with the desired outcome, showcasing the enhanced effectiveness of the protection mechanism as the Shannon Information Gain becomes more negative. Conversely, a positive value implies a gain in information, leading to a reduction in Shannon entropy and, consequently, a lower level of Internet privacy. Motahari et al. [56] provide an example of how this technique can be applied with a privacy approach. Note that measuring the Shannon Information Gain corresponds to measuring the Shannon entropy if the information system is not privacy-aware or does not provide any protection at its initial state.

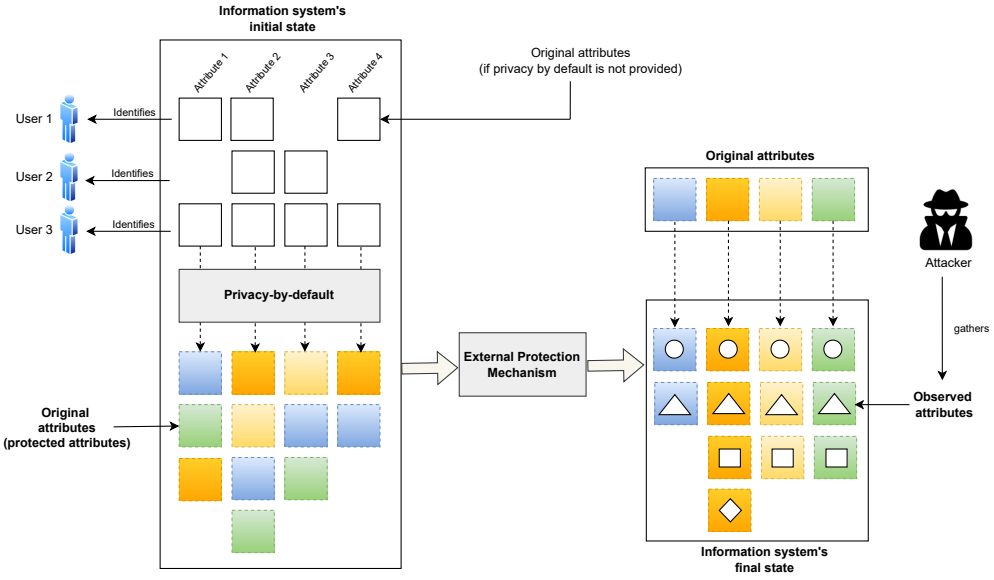

Figure 3. Application of an external protection mechanism to an information system.

Another technique is to measure the mean Mutual Information of the information system. Mutual Information is a metric based on the ratio between the conditional probability of inferring an original attribute given that a particular attribute was observed after the application of the protection mechanism, and the probability of getting the original attribute in the information system's original state. The mean value of the Mutual Information of all possible combinations between original and observed attributes is known as the mean Mutual Information metric. Since mean Mutual Information represents an increase in information, from an Internet privacy approach, it is desirable to obtain a mean Mutual Information as low as possible. Zhang et al. [58] present an example of how to apply this technique.

Although the aforementioned techniques are useful, applying Differential Privacy is the most extensively researched. This technique involves calculating the ratio between the probability of observing a subset of original attributes in the information system after applying a Differential Privacy algorithm and the probability of observing this subset in the information system's initial state. This ratio comprises the epsilon metric (ϵ-Differential Privacy) that can be interpreted as the loss of Internet privacy. When ϵ is a small positive number (e.g., $\epsilon = 0.1$), it implies a high level of Internet privacy. In contrast, when ϵ is a large positive number (e.g., $\epsilon = 1$), it denotes a low level of Internet privacy. Li et al. [59] and Jain et al. [60] present some examples of the application of this technique.

On the other hand, techniques that measure differences in information as a form of divergence are also included in the Differentiation category. One major technique involves measuring the Kullback–Leibler divergence, also known as relative entropy or I-divergence. First, the technique requires defining probability distributions for the information system's initial and final states for calculating the probabilities of each original and observed attribute. With this baseline, the relative Shannon entropy, defined as the logarithm of the ratio between the probability of an observed attribute and the probability of an original attribute, can be calculated for all possible combinations of original-observed probabilities and summed to a total divergence. While a high Kullback–Leibler divergence value suggests a high level of Internet privacy, a low Kullback–Leibler divergence value implies a low level of Internet privacy. Yang et al. [61] provide an example of this technique.

Finally, a widely used technique within this category is to measure the Sinkhorn divergence. Also known as Optimal Transport (OT) divergence, its main idea involves solving the OT optimization problem between an observed attribute and an original attribute of the information system, taking as input the statistical model defined for its initial state and its final state after applying the protection mechanism. The process is then repeated for all possible combinations of original-observed combinations, and a value of central tendency, such as the mean, is finally calculated. While low values of Sinkhorn divergence lead to less entropy and therefore more precise information involving less Internet privacy, high Sinkhorn divergence values involve more entropy and therefore a higher level of Internet privacy. A complete example of this technique can be found in [62].

Bayesian Approach: The Bayesian theory is widely used for assessing Internet privacy. Fairly similar to Differential Privacy, these techniques are based on the idea of estimating how likely it is to infer an original value from an observed one after applying a protection mechanism [63]. Therefore, as a first step, these techniques require the definition of a statistical model for the original attributes in the information system's initial state and the observed attributes in its final state. The main technique of this category involves calculating, for all possible combinations of original-observed combinations, the posterior probabilities of inferring an original attribute given that a determined attribute was observed after the implementation of the protection mechanism. Then, a metric of central tendency, such as the mean, can be calculated for the overall information system and

interpreted and linked to a level of Internet privacy. The goal is to achieve the lowest posterior probabilities for all combinations, leading to a small mean, and making it less likely to infer the original attributes from the observed ones. In [64], some examples of how to apply this technique are presented.

Another technique within this category is to conduct a Bayesian hypothesis test using both statistical models and calculating an overall Bayes factor. The Bayes factor is a metric that quantifies the evidence in favor of one hypothesis over another, based on observed data. It is defined as the ratio between the statistical evidence of getting an original attribute using the defined statistical model for the information system's final state and the statistical evidence of getting the original attribute using the statistical model defined for the information system's initial state. Therefore, an overall Bayes factor, such as the mean of the resulting Bayes factor of all possible original-observed combinations, can be computed. From an Internet privacy perspective, it is desirable to achieve an overall Bayes factor below 1, and ideally as low as possible, as this indicates weak evidence for the presence of original attributes in the data after the protection mechanism is applied. Conversely, a high overall Bayes factor reflects low effectiveness of the protection mechanism, increasing the likelihood of inferring original values from observed ones and ultimately resulting in a low level of Internet privacy. In [53], an example of the application of this technique under a privacy approach can be found.

3.1.4. Machine Learning Testing

Machine Learning has found extensive application in creating protection mechanisms aimed at enhancing Internet privacy [65,66]. Testing the effectiveness of these solutions can provide valuable insights under the PMA strategy.

Within this category, we identify two subcategories: testing regressors and testing classifiers. A common technique to test regressors is to calculate the Mean Squared Error (MSE) metric, defined as the mean of the sum of the squares of the differences between the true and predicted values of the regressor [67]. A lower MSE value indicates a more effective regressor and, usually, a higher Internet privacy. For instance, in the context of a ransomware detector system based on predicting the values of entropy due to writing patterns on disk for personal data exfiltration, a regressor with a lower MSE is considered to offer a higher level of Internet privacy compared to a regressor with a higher MSE. To calculate the Root Mean Squared Error (RMSE), defined as the root of MSE, is another equivalent technique with similar interpretation and link to Internet privacy [67]. Another technique is to calculate the Mean Absolute Error (MAE) metric. Unlike MSE and RMSE, the MAE metric involves the sum of absolute values of differences between the true and predicted values of the solution. Likewise, a low MAE value is usually interpreted and associated with a high level of Internet privacy and vice versa. Al-Ghuwairi et al. [67] offer some examples of these techniques under an Internet privacy approach.

While regression solutions can be valuable, classifiers are the primary choice for implementing protection mechanisms such as IDS solutions and detectors of privacy malware families. A prominent technique to test these classifiers is to calculate the Accuracy metric. A high Accuracy metric usually indicates a more effective protection mechanism and, therefore, a higher level of Internet privacy compared to a classifier with lower Accuracy. However, Accuracy can be misleading when dealing with imbalanced data. In scenarios involving imbalanced binary classifiers, calculating Precision and Recall comes to the forefront. While the Precision metric measures the proportion of correct positive class predictions made by the classifier, Recall determines up to what level all observations belonging to the positive class can be found. Again, a high value of these metrics is usually associated with a high level of Internet privacy. Depending on the objective of

the protection mechanisms, it might prioritize achieving a high Precision, characterized by a low false positive rate, or prioritize a high Recall, denoting a low false negative rate. In binary privacy malware classifiers, for example, the latter scenario is commonly employed, designating the privacy malware class as the positive class. For imbalanced multiclass protection mechanisms, calculating the F1-score metric, defined as the harmonic mean between Precision and Recall, can be a more reliable indicator of its effectiveness. The study presented in [68] offers an example of all these techniques.

3.1.5. Counting

The counting category is mainly used for evaluating the effectiveness of an implemented protection mechanism, and with this baseline, to define a level of Internet privacy [4]. For instance, one technique within this category is to determine the number of third-party cookies blocked by a protection mechanism on a website [69,70]. The higher the number of rejected cookies, ensuring the normal use of the website, the higher the level of Internet privacy achieved. Similarly, other techniques include to count the number of blocked trackers [71], bidders [72], and advertisements [73], as well as the number of blocked requests made by a client to third-party sites for sharing personal data while using a website [74]. The higher the number of blocked entities, the higher the level of Internet privacy achieved. Another major technique involves counting the number of queries required to identify a user on a website [75]. A high number of required queries can be interpreted as a high level of Internet privacy achieved. Identifying how well a privacy regulation aligns with normative standards and industry best practices, such as the ISO/IEC 27701 standard [76] or the NIST Privacy Framework [77], is another technique in this category. This approach helps uncover potential privacy risks by counting the number of aligned items achieved, where a high count indicates a high level of Internet privacy.

3.1.6. Similarity Measurement

From an Internet privacy perspective, similarity-based techniques aim to transform an original set of attributes from a user into a new observed set that maximizes the dissimilarity between both sets, thereby rendering it challenging to determine a user's identity [78]. Within this category, our proposed taxonomy defines three subcategories: Aggregation, Distance, and Correlation techniques.

Aggregation: The aggregation techniques aim to consolidate various original sets of attributes from different users into a single observed set, making it more challenging to infer details about a user if this observed set is gathered by an attacker [79]. A prominent technique within this category is to apply k-anonymization. As Figure 4 suggests, this technique ensures that at least k original sets of attributes can be represented by a single anonymized set of attributes [80]. A higher value of the k metric implies a higher level of Internet privacy, as it signifies an increased difficulty in inferring a specific set to identify a particular user.

Figure 5 offers an example of the application of the k-anonymization technique. Suppose that a data processor of a social network, as the case study exposed in [81], provides a service to four users and manages information such as their age, sex, salary, and zipcode. Given the sensitivity of these personal data, a protection mechanism is applied that only reveals the first digit of the age, does not disclose details about sex, indicates whether the salary exceeds a certain threshold, and omits the specific numerical details of the zipcode. These observed attributes help protect three out of the four users ($k = 3$) in case an attacker intercepts the communication, as they are not able to cover the age or salary of the fourth user. This could be interpreted as an information system that offers a medium level of

Internet privacy since it protects 75% of its users. Shin et al. [82] offer more examples of this technique.

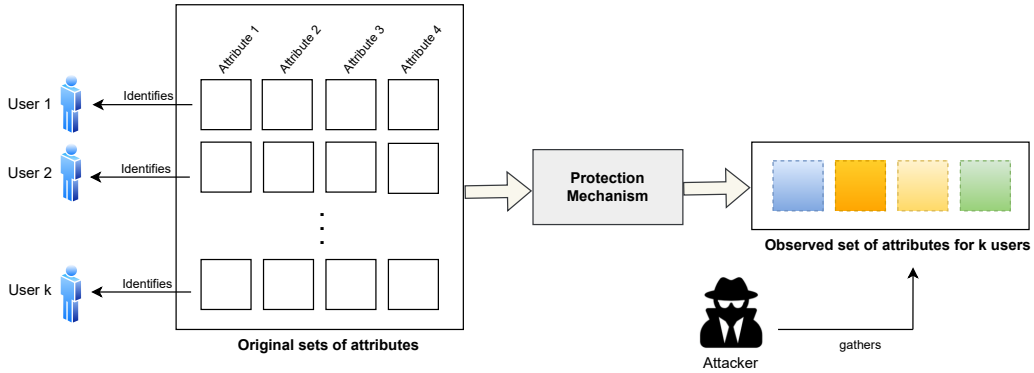

Figure 4. Application of the k-anonymization technique.

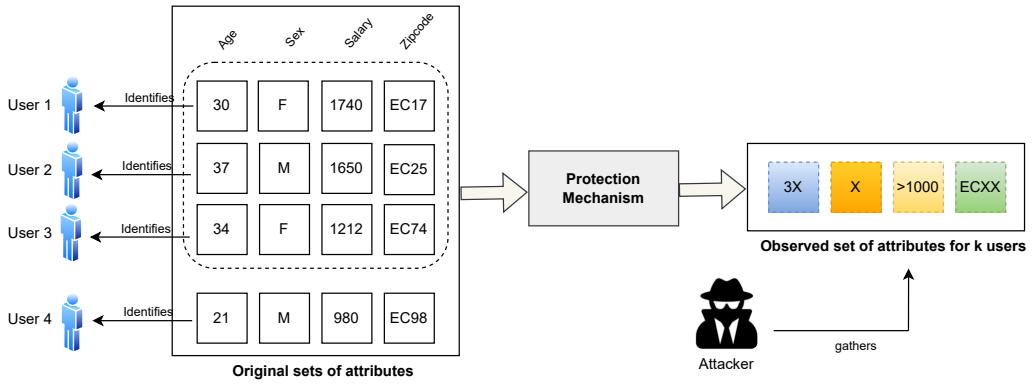

Figure 5. An example of the application of the k-anonymization technique.

Distance: As Figure 6 suggests, the techniques within this subcategory involve measuring the distance between an observed set of attributes achieved after the application of a protection mechanism (e.g., pseudonimization, encryption, noise addition, among others) and the original set of attributes of a user, where in general terms, a high distance indicates a high level of Internet privacy. A prominent technique in this category is to measure the Euclidean distance between the observed and original sets that describe a user. For this, each attribute in the original set is represented by assigning a weight value depending on its sensitivity, and similarly, a weight value is assigned to each attribute of the observed set for representing it, considering that a higher weight is assigned if the observed attribute substantially varies from the original one. Then, the Euclidean distance can be calculated as the square root of the sum of the squares of the differences in weights between corresponding elements of both sets. The overall Euclidean distance of the information system can be calculated as the mean of the Euclidean distances of each user. Lavesson et al. [83] present an example of the application of this technique.

Another technique is to calculate the Hamming distance between the original set of attributes that identifies a user and its corresponding observed set. For this, both sets are first transformed into byte arrays, and then it is calculated how many positions of their corresponding bits are different. However, both the original and observed sets must have the same byte length, which limits the use of this technique in practice. The overall Hamming distance of the information system can be calculated as the mean of the Hamming distances of each user. Kou et al. [84] and Kabwe et al. [85] offer some valuable examples of this technique.

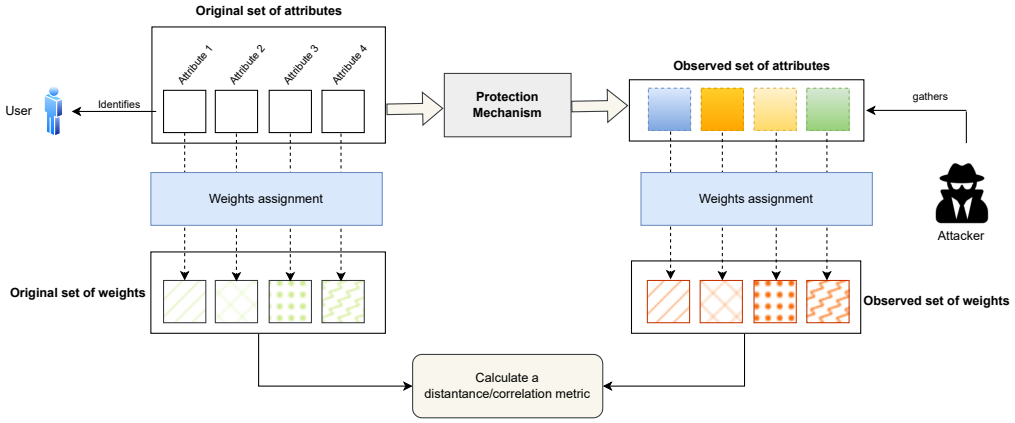

Figure 6. Application of the distance and correlation techniques.

Correlation: Among various techniques to measure correlation, the technique based on calculating the linear correlation through the Pearson correlation coefficient is one of the most commonly used to interpret and associate with a level of Internet privacy. This metric ranges between -1 and 1, and its magnitude, denoted by its absolute value, serves as a useful indicator of the similarity between the weights of the original and observed sets. Coscia et al. [86] offers an example of how the Pearson correlation coefficient is calculated and linked to an Internet privacy level. A high magnitude of the Pearson correlation coefficient suggests a high similarity and, consequently, a low level of Internet privacy. However, achieving a magnitude close to zero does not necessarily mean there is no correlation between the original and observed sets, as non-linear correlation could exist.

Finally, calculating the cosine similarity is another technique used for measuring similarity. Cosine similarity is a metric that ranges between -1 and 1 and, in the context of this research, is useful to interpret just its magnitude given by the absolute value of the dot product between the weights of both the original and observed sets, divided by the product of the magnitudes of the weights of both sets. If X denotes the original set of weights with x_i being its i-th value from L possible values, and similarly X' denotes the observed set of weights with x'_i being its i-th element from L possible values, then the magnitude of cosine similarity, denoted as $|CS|$, can be calculated using Equation (3) as follows [87]:

$$|CS| = \left| \frac{X \cdot X'}{\|X\| \, \|X'\|} \right| \tag{3}$$

where $\|X\|$ and $\|X'\|$ corresponds to the module of X and X' calculated according to Equations (4) and (5), respectively.

$$\|X\| = \sqrt{\sum_{i=1}^{L} (x_i)^2} \tag{4}$$

$$\|X'\| = \sqrt{\sum_{i=1}^{L} (x'_i)^2} \tag{5}$$

A high magnitude of cosine similarity indicates a high degree of similarity, making it easier to infer the original attributes. This, in turn, can be interpreted as a low level of Internet privacy. Figure 7 provides an example of this technique applied to a healthcare information system, taking as a baseline the case study exposed in [88]. This information system manages the blood type, sex, age, and disease of a patient. Therefore, a protection

mechanism is applied, and the original set of weights and the observed set of weights are estimated. Given that identifying a patient's specific disease could have a significant impact on the patient, greater effort is made to protect this attribute, and therefore, a higher observed weight is assigned to it compared to the other observed attributes. The vector X comprises the weight values of 3, 4, 2, 1. Vector X', on the other hand, comprises the weight values 1, 1, 20, 40. Therefore, the cosine similarity can be calculated as Equation (6) suggests, achieving a value of roughly 0.35, which could be linked to a high level of Internet privacy. More examples of this technique can be found in [87].

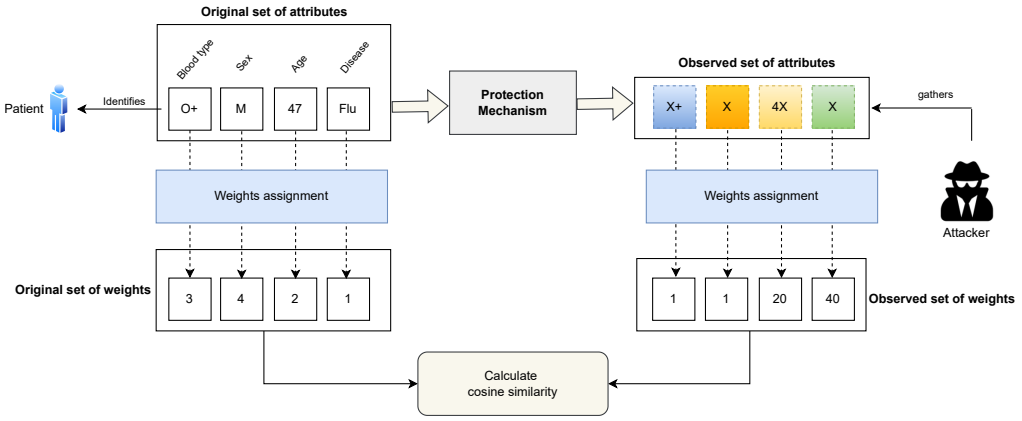

Figure 7. An example of the application of the cosine similarity technique.

$$|CS| = \left| \frac{3(1) + 4(1) + 2(20) + 1(40)}{\sqrt{3^2 + 4^2 + 2^2 + 1^2}\sqrt{1^2 + 1^2 + 20^2 + 40^2}} \right| \approx 0.35 \tag{6}$$

3.2. Levels of Assessment

Heuristic techniques and their associated metrics utilize varying scales of assessment, which are commonly categorized into four levels according to Stevens' taxonomy [89]: nominal, ordinal, interval, and ratio.

At the nominal level, data are classified into distinct categories without any intrinsic order. For example, this level can be used to evaluate whether users perceive their information as protected when using an Internet service, simply categorizing responses as "yes" or "no" [31].

When categories can be arranged in a meaningful sequence, we move into the ordinal level. A widely used method here is the Likert scale, which allows respondents to rank their perceptions, for instance, rating how secure they feel their information is on a scale from 1 (less secure) to 5 (more secure) [11]. However, while ordinal data convey order, they do not guarantee equal spacing between values, limiting the precision of quantitative interpretations [46].

This limitation is addressed at the interval level, where the distances between values are meaningful and uniform. Metrics used by machine learning testing, similarity measurement, and various counting techniques (e.g., counting the number of required queries for user identification) are within this category. Despite their enhanced precision, the interval assessment lacks an interpretation for a zero value, meaning that it is not feasible to reach or it cannot be linked to a level of Internet privacy. For example, privacy malware classifiers using regularization techniques cannot interpret a zero in their performance metric as an indication of low Internet privacy, since it may be linked to an overfitting condition in which the classifier fails in generalizing to new data. The number of queries required to identify a user cannot be zero, as at least one request is necessary to gather any information.

Similarly, a Pearson correlation coefficient equal to zero does not necessarily indicate perfect Internet privacy, as a non-linear correlation could still exist.

Finally, the ratio level incorporates all the features of interval levels but adds a meaningful zero point to its scale. Information measuring techniques usually yield ratio-level metrics, where a zero value might be linked (at least in theory) to either a total absence of Internet privacy or a state of complete Internet privacy [54,81]. The Shannon entropy is a major example, where a zero value of Shannon entropy means a total absence of Internet privacy. The correct interpretation of each scenario of analysis and how well the metric of a heuristic technique covers the problem is crucial to associate it with a level of assessment.

3.3. Levels of Automation

According to NIST SP 800-55 Vol. 1 [90], the level of automation of a solution is ascertained by evaluating the extent to which data collection, processing, and analysis are automated. These levels of automation can be classified as automatic, manual, and semi-automatic.

At the automatic level, data collection, processing, and analysis are self-operating. Automatic tools such as algorithms, bots, web crawlers, and software systems handle all aspects of the assessment process or implementation of the protection mechanism with minimal or no human intervention [67].

The manual level relies entirely on human involvement, with individuals performing data entry, processing, and analysis. This level is commonly employed in solutions that rely on human perception assessment and the interpretation of privacy policies [91]. It includes, for instance, the execution of surveys, questionnaires, and telephone calls to gather opinions, as well as manual document collection and interpretation.

The semi-automatic level combines both manual and automatic processes. While some aspects of data collection, processing, or analysis are automatic, others require human intervention or oversight. For instance, surveys and questionnaires that are executed manually but have their results analyzed using automatic statistics tools fall into this category [92].

3.4. Types of Protection Mechanisms

Based on the NIST CyberSecurity Framework [93] and the NIST Privacy Framework [77], four types of protection mechanisms have been defined: detection, prevention, response, and recovery.

Detection solutions encompass protection mechanisms designed to identify privacy threats and generate alerts or logs without taking direct action to halt an attack. This category includes alert IDS solutions, mechanisms for advertising and tracking awareness, and detectors for web browser misconfigurations and incorrect settings.

The prevention category encompasses solutions designed to safeguard privacy before an attack is executed. This includes protection mechanisms such as encryption, pseudonymization, anonymization, randomization, noise addition, obfuscation, and the detection of privacy malware based on static analysis [29].

The response category comprises solutions capable of halting an attack once it has been executed and achieves an initial impact [70]. Examples in this category include systems for blocking third-party cookies, advertising, bidders, and geolocation mechanisms once they have been executed, as well as privacy malware response systems through dynamic analysis.

Lastly, the recovery category includes protection mechanisms that respond to attacks and adapt to maintain effectiveness, while also establishing resilience and implementing

corrective actions to mitigate the impact of attacks [94]. Examples in this category are self-adjusting anonymization, advanced privacy malware classifiers, and sanitization solutions provided by advanced privacy protection systems [95].

3.5. Broad Domains

Classifying a solution into a specific domain can be challenging due to the overlapping nature of multiple existing domains. However, the MITRE ATT&CK Framework offers a practical approach for categorizing solutions into the broad domains: Enterprise, Mobile, and Industrial Control Systems [96].

The Enterprise broad domain encompasses the systems, networks, and infrastructure of organizations, businesses, and individuals consuming Internet services through a web browser. This category includes solutions in fields such as Web, Cloud Computing (CC), and Fog Computing (FC), among others.

The Mobile broad domain focuses on solutions that enable ubiquitous connectivity. This includes not only smartphones and Personal Digital Assistants (PDA) widely analyzed in the Mobile Computing (MC) domain, but also the paramount fields of the IoT, Internet of Vehicles (IoV), and Vehicle Ad hoc Networks (VANETs), among others.

The Industrial Control Systems broad domain includes solutions for managing and controlling industrial processes, machinery, appliances, and equipment across diverse sectors such as manufacturing, energy, utilities, and transportation. This category includes the Industrial Internet of Things (IIoT) and Supervisory Control and Data Acquisition (SCADA) systems, among others.

4. Methodology

In order to carry out the review, we followed a methodological framework consisting of the stages illustrated in Figure 8. The methodology began with the planning phase, where we established the primary goal of the review and defined the Research Questions (RQs). During this stage, we also developed the review protocol, which involved defining the paper selection strategy, setting inclusion/exclusion criteria, creating a classification scheme, defining an extraction process, determining visualization methods, and addressing potential threats to validity.

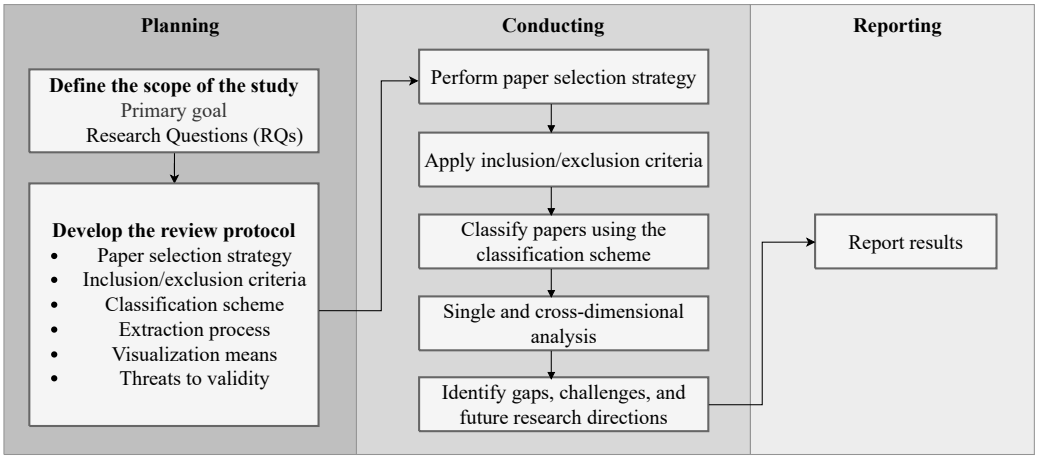

Figure 8. Methodology.

The second stage involved conducting the review. In this stage, we implemented the previously defined paper selection strategy to gather relevant papers and applied the inclusion/exclusion criteria. We classified the selected papers using the predefined classification scheme and conducted both a single and cross-dimensional analysis to identify gaps in the literature as well as to analyze challenges and foretell future research directions.

Finally, in the third stage, we formulated responses to the RQs, allowing us to draw our final research conclusion and report the results we obtained from the review.

4.1. Scope of the Study

In our methodology, the first stage entailed defining the scope of the study by establishing the primary goal of the review and formulating the associated RQs. Our primary goal was to identify the main heuristic techniques used for assessing Internet privacy and their related protection mechanisms. The defined RQs were the following:

RQ1. What are the main heuristic techniques and their associated metrics used for assessing Internet privacy?

Motivation: Internet Privacy is a subjective term open to various interpretations. To assess Internet privacy effectively, each heuristic technique employs a metric that supports subsequent analysis, interpretation, and linkage to a level of Internet privacy. Consequently, a variety of heuristic techniques relying on diverse metrics can be applied.

RQ2. At what level is Internet privacy assessed through the main heuristic techniques?

Motivation: The various heuristic techniques and their associated metrics used for assessing Internet privacy have limitations when applied in practical scenarios. Some metrics offer a criterion of order and allow performing comparisons, while others require additional criteria to establish a connection with Internet privacy.

RQ3. At what level of automation is Internet privacy assessed through the main heuristic techniques?

Motivation: Assessing Internet privacy involves collecting and manipulating information using various tools and solutions. Depending on the technologies employed and their level of automation, each heuristic technique could be executed either automatically, semi-automatically, or manually.

RQ4. What types of protection mechanisms have been applied in order to address the lack of privacy when using services over the Internet?

Motivation: Numerous services are provided to users over the Internet, and the lack of privacy is a significant concern. Therefore, it is crucial to understand the types of protection mechanisms that have been implemented to address the absence of privacy when using Internet services.

RQ5. In what broad domains are heuristic techniques employed for the assessment of Internet privacy?

Motivation: The Internet encompasses a wide range of applications and services, not solely limited to the web scenario. It also extends to various broad domains, including Mobile and Industrial Control Systems. Therefore, this review has been carried out in a domain-agnostic manner in order to link each contribution to its general broad domain.

4.2. Paper Selection Strategy

After establishing the primary goal and the RQs, a paper selection strategy was carefully devised. As sources of information, all papers were obtained from the Scopus, IEEE Xplore, ACM Digital Library, and SpringerLink databases. The reason for choosing these databases is their well-established relevance in the fields of Information Security and Privacy, with numerous high-impact scientific papers. Additionally, these databases enabled us to identify prominent journals and conferences, including those affiliated with ACM, IEEE, Elsevier, among others.

With this baseline, we identified four keywords associated with our primary goal, and then we defined related terms such as synonyms and associated concepts, referring to the IEEE Thesaurus guide for advisory [97]. Table 3 details these keywords and related terms.

Table 3. Keywords and related terms.

Keyword	Related Terms
Assessment	Measurement, Metric
Internet privacy	Online privacy
Heuristic technique	Method, Methodology, Procedure, Strategy
Protection mechanism	Control, Countermeasure, Defense mechanism, Privacy-enhancing technologies (PETs)

Building upon this foundation, and acknowledging the necessity to identify papers that propose heuristic techniques for assessing Internet privacy and/or protection mechanisms to achieve our primary goal, we formulated the general search string depicted in Figure 9.

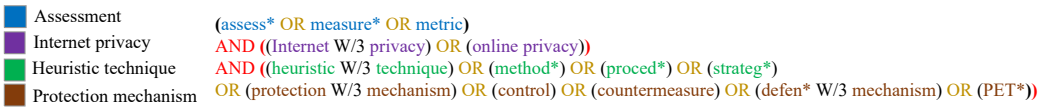

Figure 9. General search string. Each related term forming the search string is associated with one of the four keywords through its corresponding color. The symbol * denotes a wildcard, indicating that additional characters may follow the initial word.

4.3. Inclusion and Exclusion Criteria

Automatic and manual inclusion and exclusion criteria were systematically applied to filter the retrieved papers obtained from the four databases. The process began with the application of the automatic inclusion and exclusion criteria. Subsequently, manual inclusion and exclusion criteria were applied in a screening phase.

Table 4 provides details of the automatic inclusion criteria that were used to narrow down the number of papers. The corresponding exclusion criteria were formulated based on the opposite conditions of the inclusion criteria. The search was restricted to include only final conference or journal papers written in English within the field of Computer Science.

Table 4. Automatic inclusion criteria.

Filter	Inclusion Criteria
Language	English
Document type	Conference Paper or Journal
Research field	Computer Science
Publication stage	Final

On the other hand, manual inclusion and exclusion criteria were defined to carry out the screening phase. Table 5 outlines the manual inclusion criteria of this review. A **YES** response to all these criteria was mandatory for paper inclusion, while any **NO** response resulted in paper exclusion. The exclusion criteria were established as the inverse of the inclusion criteria.

Table 5. Manual inclusion criteria.

Filter	Inclusion Criteria
Type of contribution	Is the paper a primary contribution?
Internet privacy related	Does the paper present a solution related to Internet privacy or one that could be adapted to Internet privacy?
Paper contribution	Is the paper contribution related to a technique and/or a protection mechanism able to assess or protect Internet privacy?
Contribution proof	Does the paper provide a proof that helps corroborate the proposed solution?

4.4. Screening Phase

The screening phase was divided into three sequential stages: title screening, abstract screening, and full-paper screening. At each stage, manual inclusion and exclusion criteria were applied and, based on the results obtained, one of the following marks was assigned to each paper as follows:

[I]ncluded. This mark was assigned if the paper fulfilled all the manual inclusion criteria and none of the manual exclusion criteria. However, as a unique exception to this rule, secondary contributions that complied with all the rest of the manual inclusion criteria, during the title and abstract screening, were also marked as Included for allowing them to reach the full-paper screening stage.

[S]econdary. This mark was used during the full-paper screening for all secondary papers that complied with all the rest of the manual inclusion criteria. Since the paper did not meet the first manual inclusion criterion, it was not finally considered. However, reading and analyzing the paper allowed us to obtain knowledge and a general map of related work.

[E]xcluded. The paper failed to meet at least one manual inclusion criterion or met at least one exclusion criterion. Therefore, the paper was discarded.

[U]nclear A dilemma arising from the paper analysis was thoroughly discussed in the next screening stage, unless this comprised the final full paper screening step, where a definitive decision had to be made.

During the title screening, all papers with Excluded marks were discarded. The remaining papers then underwent abstract screening, during which the papers with Excluded marks were once again removed. The papers that passed these two screening stages served as the input for the full-paper screening. In this final step, the Unclear mark was not used, resulting in papers being finally marked as Included, Secondary, or Excluded.

To minimize potential bias, three coders conducted the complete screening process, which consisted of both a pilot phase and a main phase. The pilot phase aimed to establish consensus within the research group. In this phase, a random sample of 35 papers was selected after applying automatic inclusion and exclusion criteria. During the pilot title screening, each paper was assigned one of the following marks: Included, Excluded, or Unclear. Secondary papers were marked as Included if they met all the remaining inclusion criteria. After marking all available papers, the Krippendorff's Alpha coefficient was calculated. If this coefficient did not reach a minimum value of 0.85, multiple meetings were held to align opinions in the next iteration and, if necessary, additional papers were also reviewed to achieve consensus. All Included and Unclear papers resulting from this process entered the pilot abstract screening, where a similar procedure was carried out. After achieving a Krippendorff's Alpha coefficient of at least 0.85, the papers designated as Included and Unclear advanced to the pilot full-paper screening stage, which followed the same process but without considering the Unclear mark. After reaching a Krippendorff's Alpha coefficient of at least 0.85, all papers were finally marked with a unified grading criterion defined by the coders.

Building on this foundation, the main screening phase detailed in Figure 10 was conducted. Initially, all the papers that were not assessed during the pilot were selected as input for the main screening phase. Subsequently, each paper underwent the established screening procedure. Any discrepancies or disagreements that arose were thoroughly reviewed and resolved during multiple work meetings, where a final decision was made.

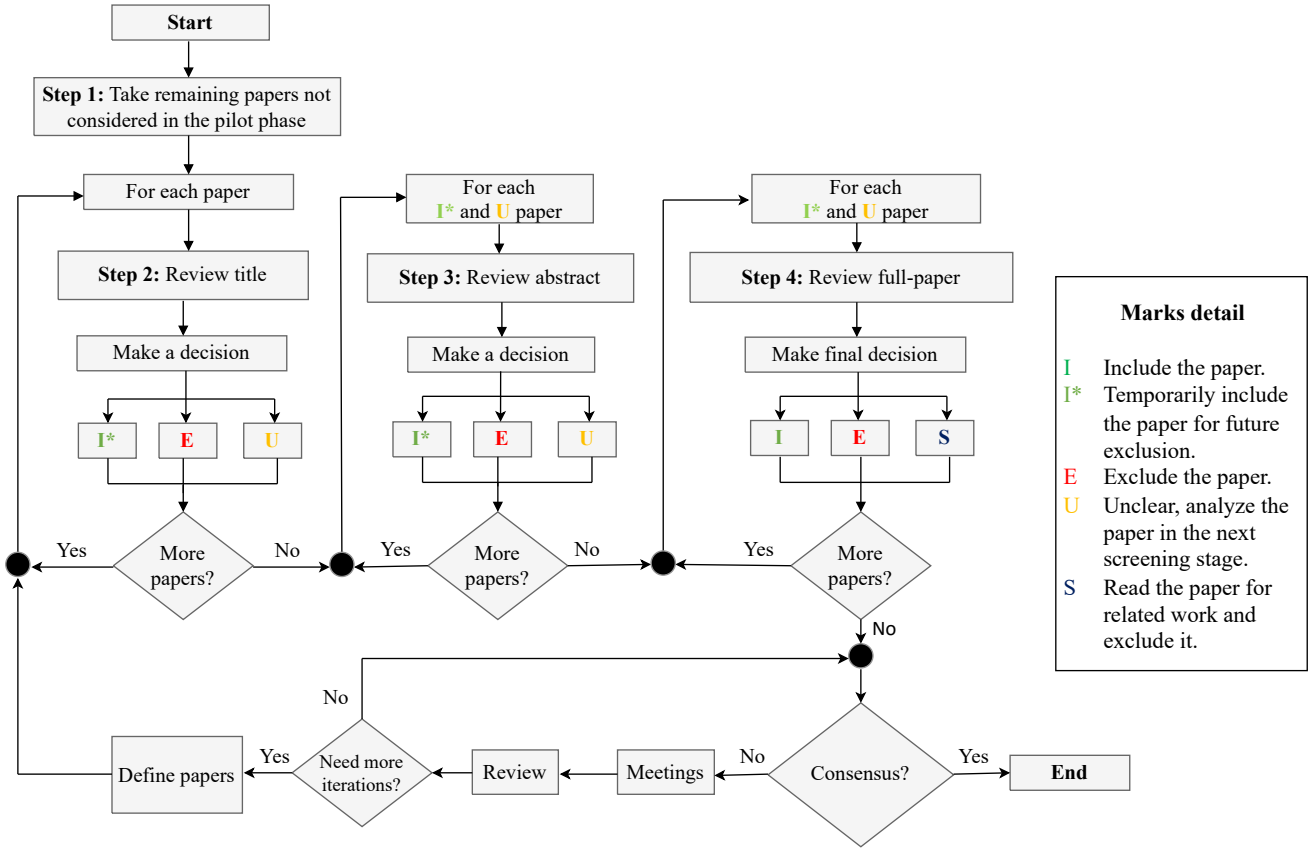

Figure 10. Main screening phase.

4.5. Classification Scheme

Figure 11 details our classification scheme, which is conformed by: (a) Heuristic Techniques Categories, housing our proposed taxonomy for a comprehensive classification of existing heuristic techniques for assessing Internet Privacy; (b) Levels of Assessment, encompassing nominal, ordinal, interval, and ratio levels; (c) Levels of Automation, inclusive of the NIST SP 800-55 Vol. 1 [90] automatic, semi-automatic, and manual levels; (d) Types of Protection Mechanisms, where are found detection, prevention, response, and recovery mechanisms as defined by the NIST Cybersecurity Framework and NIST Privacy Framework; and (d) Broad Domains, which considers the general domains of Enterprise, Mobile, and Industrial Control Systems outlined by the MITRE ATT&CK Framework.

4.6. Coding Procedure

The coding procedure facilitated the systematic location, arrangement, and categorization of all papers marked as Included (resulting from the screening phase) into the various categories outlined in our proposed classification scheme. This methodical organization of knowledge facilitated a robust identification of gaps, challenges, and promising directions for future research.

To minimize subjectivity and bias, our process incorporated both a pilot coding phase and a main coding phase, involving the evaluation of three coders at each stage. The pilot coding phase started with a random selection of ten papers, with coders evaluating them until a Krippendorff's Alpha coefficient of at least 0.85 was achieved. In cases where this threshold was not reached, collaborative working sessions were convened to refine and harmonize the criteria for a new iteration of the pilot coding phase on the selected ten papers. Our methodology also allowed for the inclusion of additional papers in the pilot if

necessary to reach a consensus. Consequently, the pilot coding phase concluded once the Krippendorff's Alpha coefficient reached a minimum of 0.85.

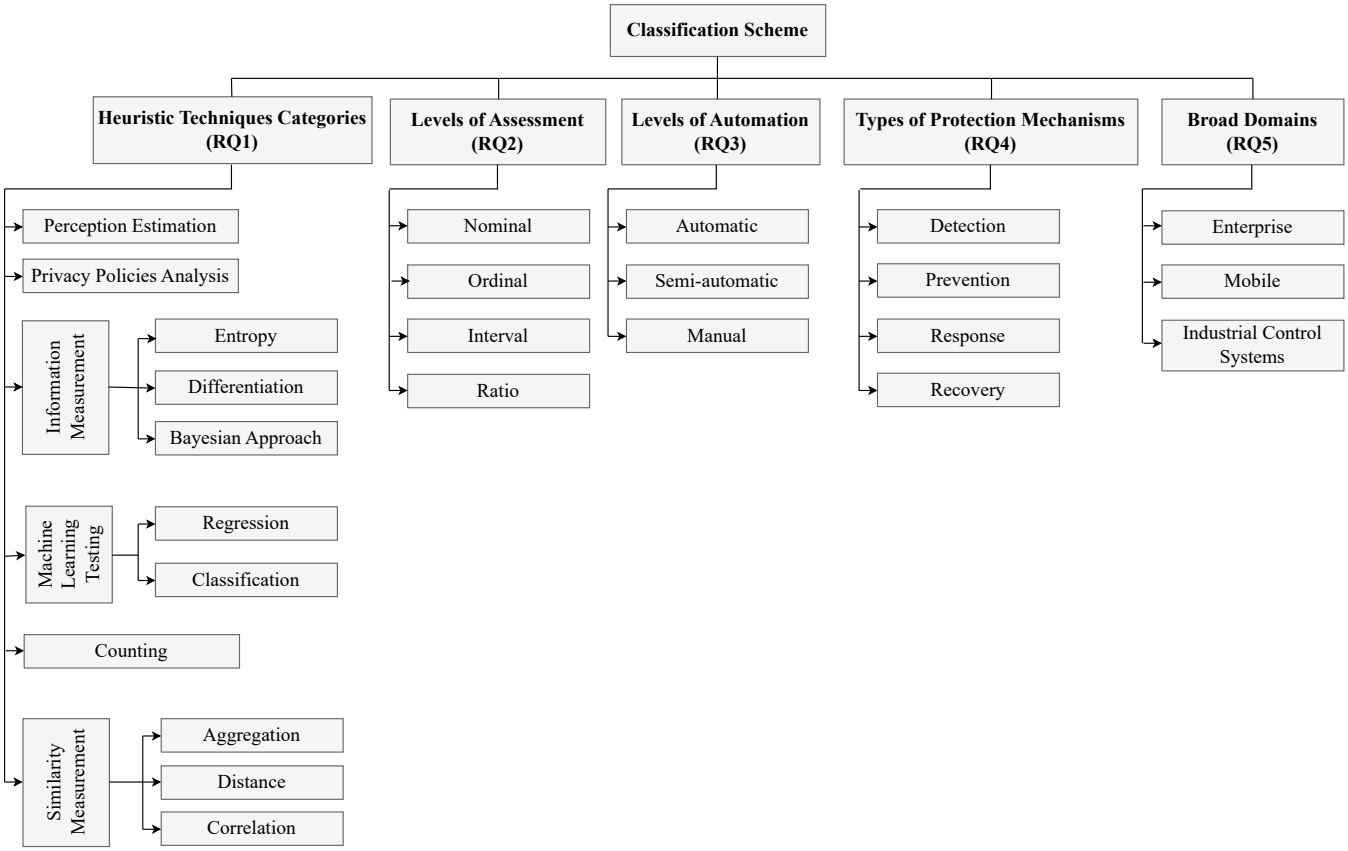

Figure 11. Classification scheme.

After harmonizing the evaluation criteria among the coders, the main coding phase was initiated, encompassing the remaining papers not covered in the pilot coding phase. Any disparities or discrepancies were thoroughly examined during collaborative work sessions, allowing for the presentation of diverse viewpoints and the ultimate resolution of coding decisions for the papers.

4.7. Paper Selection and Screening Results

We were able to retrieve a total of 5114 papers from the four databases, which were subsequently narrowed down to 934 papers through the application of automatic inclusion and exclusion criteria and duplication removal. Out of these 934 papers, 35 were randomly chosen for the pilot screening phase.

For the pilot title screening, two iterations were required to attain a final Krippendorff's Alpha coefficient of 0.902. Similarly, in the pilot abstract screening, two iterations resulted in a final Krippendorff's Alpha coefficient of 0.851. Lastly, the pilot full paper screening involved three iterations, culminating in a Krippendorff's Alpha coefficient of 0.869. With this baseline, the main screening phase was conducted. As the overall result of the screening process, of the initial 934 papers, 714 remained after the title screening. Following the abstract screening, 329 papers were retained, and after the full-paper screening, a total of 160 papers were finally included for the coding phase. Figure 12 presents the PRISMA flow chart of our review detailing the results of this process.

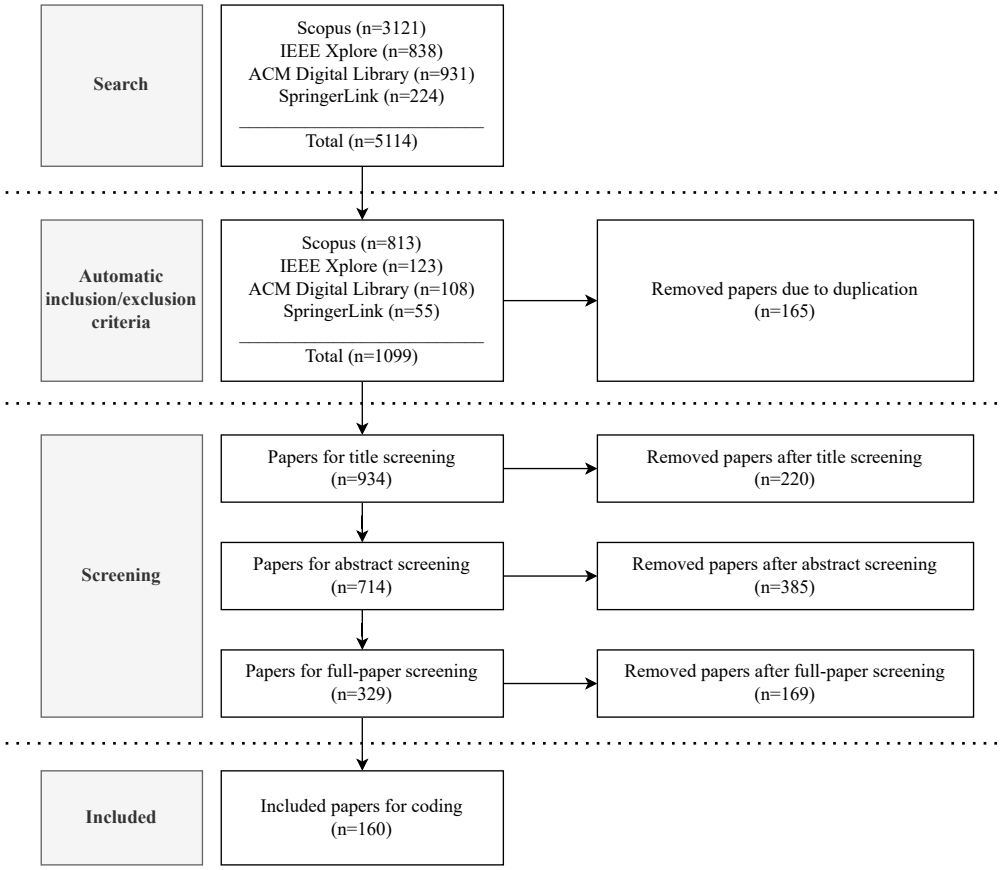

Figure 12. PRISMA flow chart.

4.8. Pilot Coding Phase

From the pool of 160 papers identified during the screening phase, we randomly selected 10 papers to undergo the pilot coding phase. The process involved a total of five iterations, culminating in the attainment of a conclusive Krippendorff's Alpha coefficient of 0.875. After reaching consensus among team members and standardizing the criteria, the main coding phase seamlessly unfolded. The comprehensive review of the papers enabled us to address the defined RQs, leading to the results presented and analyzed in Section 5.

5. Results and Analysis

In this section, we present and analyze the outcomes derived from our proposed methodology. The findings are based on a comprehensive analysis and coding of all 160 papers defined through the screening phase. Figure 13 illustrates the number of papers published each year. Notably, the results reveal a significant increase in research over the last three years, driven by the proposal, enactment or entry into effect of privacy regulations worldwide, such as the American California Privacy Rights Act (CPRA), in effect from 2023 [98], and the Canadian Consumer Privacy Protection Act (CPPA), proposed in 2022 [99].

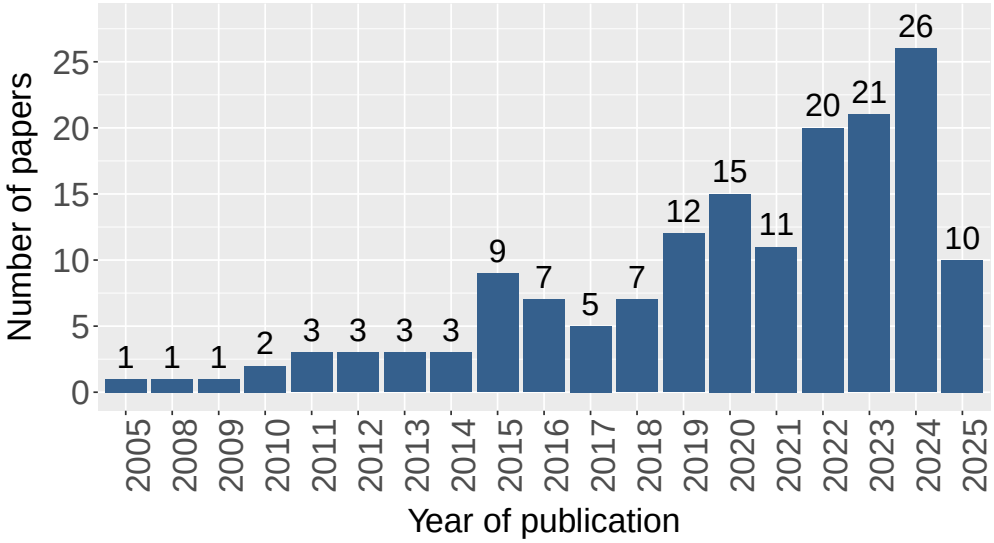

Figure 13. Number of papers by year of publication.

5.1. Heuristic Techniques

In this subsection, we answer our RQ1: What are the main heuristic techniques and their associated metrics used for assessing Internet privacy?. Table 6 summarizes the main heuristic techniques within each category and their associated metrics, as well as the techniques' connection to a level of Internet privacy. Each heuristic technique involves measuring its corresponding metric for a subsequent analysis, interpretation, and association with a level of Internet privacy. The table also offers a comparison between heuristic techniques, exposing their advantages and disadvantages.

Figure 14 summarizes the total number of papers by heuristic technique category. Overall, most heuristic techniques predominantly rely on information measurement methods, followed by privacy policies analysis and counting techniques. Perception estimation and similarity measurement have also been widely used to assess Internet privacy. However, identifying new ways to analyze and link the performance of machine learning solutions to levels of Internet privacy needs more research. Table 7 details the papers within each heuristic technique category.

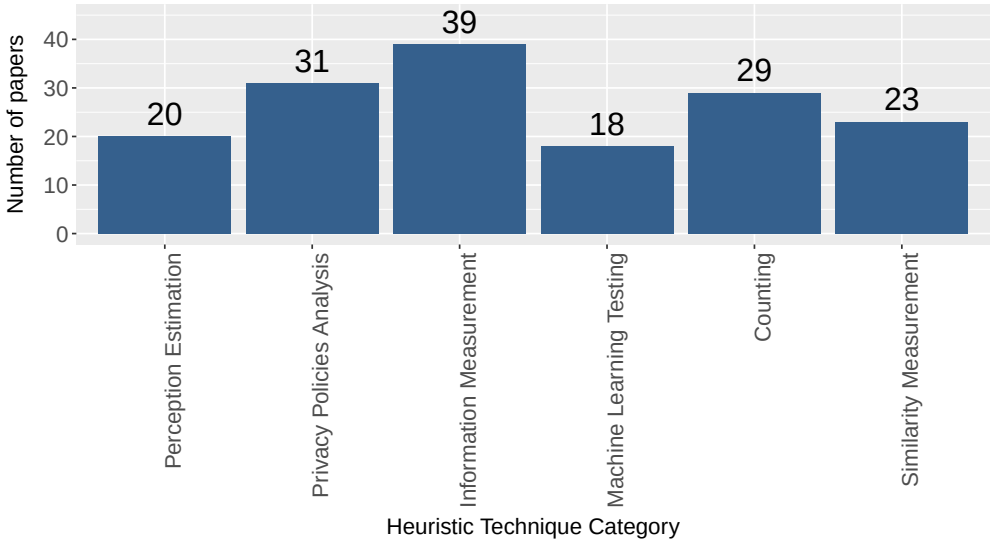

Figure 14. Number of papers by heuristic technique category.

Table 6. A comparison of the main heuristic techniques for assessing Internet privacy.

N°	Heuristic Technique Category	Metric	Link to Privacy	Advantages	Disadvantages
1	Perception Estimation	Likert scale [31]	Depends on the context (e.g., a high number of answers choosing option 3 in a question like "How much protected do I feel is my personal data being 1—low, 2—medium, 3—high?" involves a high Internet privacy).	-Easy to interpret.	-Highly subjective and based on opinions.
2	Privacy Policies Analysis	Flesch Reading Ease [49]	A high metric involves a high Internet privacy.	-Easy to calculate. -Simple mathematical formulation.	-English Language oriented metric. -Does not determine if statements in the privacy policy are implemented in practice.
3	Information Measurement-Entropy	Shannon Entropy [54]	A high metric involves a high Internet privacy.	-Easy to calculate.	-Reaches its maximum value just on equiprobable attributes.
4	Information Measurement-Differentiation	Shannon Information Gain [100]	A high metric involves a low Internet privacy.	-Simple to calculate.	-Requires two calculations of entropy.
5	Information Measurement-Differentiation	Mean Mutual Information [15]	A high metric involves a low Internet privacy.	-Low processing cost.	-Can consider already analyzed and redundant attributes as part of its calculus.
6	Information Measurement-Differentiation	ϵ-Differential Privacy [14]	A small ϵ value involves a high Internet privacy.	-Widely researched. -Offers strong resistance to re-identification attacks.	-Requires implementing an algorithm in software to be applied. -Resource expensive.
7	Information Measurement-Differentiation	Kullback–Leibler divergence [61]	A high metric involves a high Internet privacy.	-Its use demonstrates good results. -Is the base for multiple derived metrics (e.g., Jensen-Shannon Divergence).	-Does not meet the triangular inequality criterion. -Low quality representation in latent distributions. -It is a non-symmetric metric.
8	Information Measurement-Differentiation	Sinkhorn divergence [62]	A high metric involves a high Internet privacy.	-Based on optimization methods. -Offers a more realistic metric.	-Complex to calculate. -Usually time-consuming.

Table 6. *Cont.*

N°	Heuristic Technique Category	Metric	Link to Privacy	Advantages	Disadvantages
9	Information Measurement-Bayesian Approach	Posterior probability [101]	A high probability involves a low level of Internet privacy.	-Simple to calculate once conditional probabilities are defined.	-Requires defining statistical models for the information system's initial and final states. -The principles of independence of events required for problem simplification are usually difficult to interpret.
10	Information Measurement-Bayesian Approach	Overall Bayes factor [18]	A high metric involves a high Internet privacy.	-Offers a more realistic approach than binary hypothesis tests.	-Requires defining statistical models for the information system's initial and final states.
11	Machine Learning Testing-Regression	MSE [67]	Usually, a high metric involves a low Internet privacy.	-Easy to calculate. -Not computationally demanding.	-Difficult to interpret. -Sensible to central tendency bias.
12	Machine Learning Testing-Classification	Accuracy [102]	Usually, a high metric involves a high Internet privacy.	-Good for binary classification problems.	-Susceptible to unbalanced data.
13	Machine Learning Testing-Classification	F1-score [102]	Usually, a high metric involves a high Internet privacy.	-Good for multiclass classification problems. -Good for unbalanced data.	-Sensible to variations in Precision and Recall.
14	Counting	Various (e.g., Number of third-party cookies blocked [5])	Depends on the context, e.g., a high number of third-party cookies blocked involves a high Internet privacy.	-Easy to interpret.	-Susceptible to contexts of analysis. -Depending on the context, it can make it difficult to take a representative sample in real-world conditions.
15	Similarity Measurement-Aggregation	k-anonymity [103]	A high k value involves a high Internet privacy.	-Easy to calculate and interpret.	-If broken, all the k sets of original attributes are in potential danger. -Not suitable for multidimensional problems.

Table 6. *Cont.*

N°	Heuristic Technique Category	Metric	Link to Privacy	Advantages	Disadvantages
16	Similarity Measurement-Distance	Euclidean distance [104]	A high metric involves a high Internet privacy.	-Easy to calculate.	-Defining the weights of attributes requires additional analysis and interpretation. -Does not capture semantic differences.
17	Similarity Measurement-Distance	Hamming distance [105]	A high metric involves a high Internet privacy.	-Fast in software due to binary operations.	-Requires that both sets of attributes have the same length in bytes.
18	Similarity Measurement-Correlation	Pearson correlation coefficient [86]	A high metric involves a low Internet privacy.	-It does not depend on the units of measurement.	-Only detects linear correlation. -Sensitive to outliers.
19	Similarity Measurement-Correlation	Cosine similarity [87]	A high metric involves a low Internet privacy.	-Works well with sparse and high-dimensional data. -Computationally efficient.	-It primarily considers the orientation rather than the magnitude of the attribute representations.

Table 7. Detail of papers by heuristic technique category.

Heuristic Technique Category		Count	Papers
Perception Estimation		20	[31,32,46,47,91,106–120]
Privacy Policies Analysis		31	[9,11,48–52,121–144]
Information Measurement	Entropy	7	[54–56,145–148]
	Differentiation	22	[10,14,15,57–62,81,88,94,100,149–157]
	Bayesian Approach	10	[18,53,63,64,101,158–162]
Machine Learning Testing	Regression	1	[67]
	Classification	17	[6,65,66,68,95,163–172]
Counting		29	[1,3–5,7,8,12,69–75,92,173–186]
Similarity Measurement	Aggregation	7	[17,79,80,82,103,187,188]
	Distance	11	[13,83–85,104,105,189–193]
	Correlation	5	[78,86,87,194,195]

5.2. Levels of Assessment

This subsection provides an answer to our RQ2: At what level is Internet privacy assessed through the main heuristic techniques? Figure 15 exposes a heatmap with the cross-dimensional results of number of papers classified by heuristic technique category and level of assessment, level of automation, and broad domain. The figure demonstrates that Internet privacy is assessed through the four levels of Stevens' taxonomy.

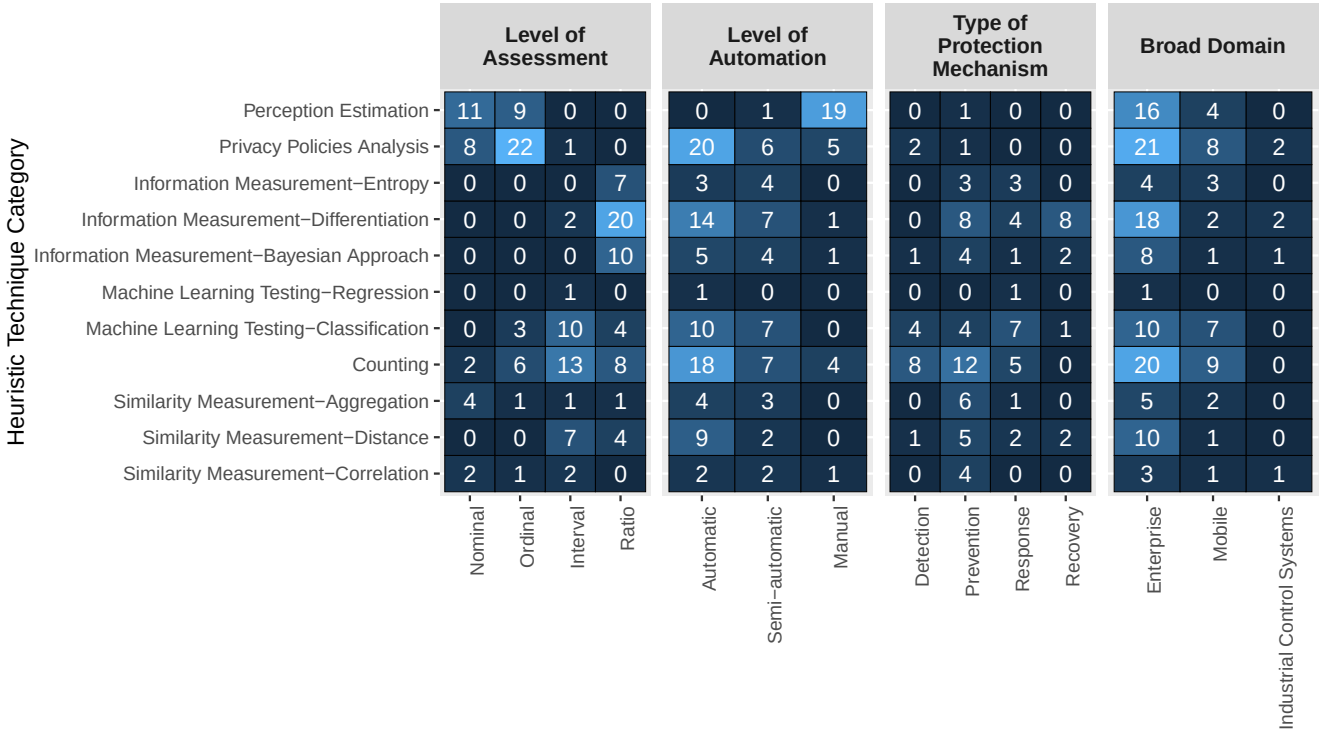

Figure 15. Number of papers by heuristic technique category and level of assessment, level of automation, type of protection mechanism, and broad domain. The dark blue cells indicate a small number of papers, while the light blue cells indicate a high number.

However, most research is focused on assessing Internet privacy at the ratio level, driven by information measuring techniques. While privacy policies are mostly assessed at the ordinal level, the techniques based on measuring the achieved difference in information predominantly employ the ratio scale. The interval level, in contrast, is mostly used by

solutions based on counting, machine learning classifiers, and similarity measurement techniques through the calculus of a distance metric. Table 8 details the advantages and disadvantages of each level of assessment and presents the main metrics associated with each level. Table 9 provides the details of papers by level of assessment. Notably, one-third of the existing papers propose solutions at the ratio scale level.

Table 8. A comparison of levels of assessment.

Level of Assessment	Main Metrics	Advantages	Disadvantages
Nominal	-Yes/no questions [31].	-Easy to estimate.	-Analysis limited due to the absence of an order criterion.
Ordinal	-Likert scale [124]. -Flesch Reading Ease [49].	-Allows to determine if a solution is better than others.	-Does not allow for making comparisons in the sense of magnitude, that is, to define how much better a solution is than order.
Interval	-k metric [80]. -Number of required queries [75]. -Accuracy of classifiers with no regularization methods [102].	-Allows for making comparisons between solutions in the sense of magnitude.	-Lack of interpretation for the zero value.
Ratio	-Number of disclosed personal data [176]. -Number of blocked third-party cookies [5]. -Bayes factor [18]. -ϵ-Differential Privacy [14].	-Allows more precision when making comparisons. -Defines a criterion for the zero value of its scale.	-Usually mathematically more complex. -Usually requires the definition of statistical models.

Table 9. Detail of papers by level of assessment.

Level of Assessment	Count	Papers
Nominal	27	[11,12,17,31,50,52,78,86,103,106–109,111–114,119,120,125,127,132,137,141,179,187,188]
Ordinal	42	[9,32,46,47,49,51,66,71,72,82,87,91,110,115–118,121–124,126,128–131,133–136,138–140,142–144,168,170,178,181–183]
Interval	37	[4,7,10,13,16,48,67–70,73–75,80,81,83–85,92,95,102,163–167,169,177,180,184–186,189,190,193–195]
Ratio	54	[1,3,5,6,8,14,15,18,53–65,79,88,94,100,101,104,105,145–162,171–176,191,192]

5.3. Levels of Automation

This subsection addresses our RQ3: At what level of automation is Internet privacy assessed through the main heuristic techniques? Figure 15 shows that Internet privacy is assessed through automatic, semi-automatic, and manual approaches. While perception estimation is still conducted using manual tools such as surveys and questionnaires, automatic tools are primarily employed by heuristic techniques such as privacy policies analysis, counting, information measurement by differentiation, machine learning classifiers, and similarity measurements using distance metrics. Semi-automatic tools, on the other hand, are used by certain counting techniques that implement automatic processes to retrieve information for subsequent manual analysis (e.g., analyzing information from web crawlers), as well as by information measurement techniques based on entropy, which require analyzing the frequency and probabilities of attributes' appearance within an information system. Table 10 provides a comparison of levels of automation used to assess Internet privacy, along with some examples of the most common tools of each category. Table 11 details the papers for each level of automation.

Table 10. A comparison of levels of automation for assessing Internet privacy.

Level of Automation	Main Tools	Advantages	Disadvantages
Automatic	-Algorithms [14]. -Web crawlers with reports functionality [135]. -Platforms for compliance analysis [51].	-Provides results in less time. -Offers a more granular interpretation of results [196,197].	-Implementation of platforms could be expensive and/or time-consuming. -Automatic tools may be restricted in certain environments, limiting their ability to gather sensitive information that a human could obtain.
Semi-automatic	-Manual analysis of websites' information retrieved from web crawlers [60]. -Manual definition of frequencies and probabilities of attributes' appearance for automatic entropy calculation [151].	-Combines automated efficiency with human judgment.	-Offers less scalability than automatic tools. -The process can become slow when dealing with large volumes of data.
Manual	-Surveys [31]. -Questionnaires [106]. -Telephone calls [91]. -Manual analysis of information disclosed in websites and portals [12].	-Easy to perform. -Access to information in environments with protection against automatic tools.	-Prone to errors. -Time-consuming.

Table 11. Detail of papers by level of automation.

Level of Automation	Count	Papers
Automatic	86	[3–5,8–10,14–16,18,48,51,53,56–58,61,65–72,74,79,80,83–85,87,94,101–104,121,122,125,126,128–130,132–135,138–143,146,147,150,152–157,159,161,163,167,168,170,171,173–178,182,183,186,187,189–194]
Semi-automatic	43	[1,6,7,13,17,49,54,55,59,60,62–64,73,75,78,81,82,86,88,92,95,105,117,124,127,131,136,144,145,148,149,151,158,162,164–166,169,172,180,185,188]
Manual	31	[11,12,31,32,46,47,50,52,91,100,106–116,118–120,123,137,160,179,181,184,195]

5.4. Types of Protection Mechanisms

In this subsection, we will provide an answer to our RQ4: What types of protection mechanisms have been applied in order to address the lack of privacy when using services over the Internet? From the 160 papers included for coding, just 101 propose a protection mechanism to interpret and link their effectiveness to a level of Internet privacy (using the PMA approach as done, for example, by machine learning classifiers for ransowmare detection, IDS, and IPS solutions) or to exemplify the use of a heuristic technique to associate its results with the level of effectiveness achieved by the protection mechanism (e.g., applying the Bayesian approach techniques to determine the effectiveness of an anonymization and noise addition protection mechanism). Figure 15 suggests that Internet privacy is protected through the four categories of protection mechanisms derived from the NIST CyberSecurity Framework [93] and the NIST Privacy Framework [77]. However, counting techniques and techniques based on information measurement by differentiation are mostly applied in preventive solutions. Particularly, this last heuristic technique category is predominant for determining the level of effectiveness of recovery solutions. In contrast, machine learning classifiers are widely used for the detection of compliance solutions, as well as for prevention and response to privacy malware. Table 12 offers a comparison of the main protection mechanisms for enhancing Internet privacy detailing their advantages and disadvantages.

Table 12. A comparison of the main protection mechanisms for enhancing Internet privacy.

N°	Type of Protection Mechanism	Protection Mechanism	Advantages	Disadvantages
1	Detection	Alert IDS [1]	-Usually comprises standalone solutions. -Easy to deploy.	-Limits its actions to just generating logs or alerts with no actions. -Allows the threats to cause an initial impact. -Fine-tuning of signatures required.
2	Prevention	Encryption [3]	-Significantly mitigates risks of unauthorized access to personal data.	-Key management is complex and critical. -Performance overhead for strong encryption algorithms [198]. -Encrypted data is usually unusable (an exception is Homomorphic encryption that allows computations to be performed without needing to decrypt data first [66]).
3	Prevention	Pseudonymization [160]	-Enables data utility for analytics while protecting identity.	-Allows re-identification through auxiliary data. -Requires secure management of mapping keys. -Not suitable for high-sensitivity datasets without additional countermeasures.
4	Prevention	Anonymization [14]	-Supports compliance with strict privacy regulations. -Enables data sharing without violating privacy regulations.	-No standardized anonymization method for all scenarios.
5	Prevention	Randomization [17]	-Simple to implement in many statistical contexts. -Scales well for large datasets.	-Susceptible to brute force attacks.
6	Prevention	Noise addition [104]	-Simple and mathematically sound for privacy guarantees.	-Requires careful noise calibration to avoid excessive distortion.
7	Prevention	Obfuscation [157]	-Low implementation complexity for many cases. -Low computational cost.	-Can be bypassed through advanced analysis or reverse engineering.
8	Prevention	Masking [87]	-Easy to implement by systems allowing privacy-by-default.	-Extremely susceptible to brute force attacks.
9	Prevention	Contextual integrity [171]	-Defines authorized privacy flows for protection. -Supports adaptive and flexible implementation of privacy policies.	-Requires deep understanding of varied contexts. -Limited support for automation and tools.
10	Prevention	DLP [158]	-Enforces data handling policies. -Monitors data across endpoints, network, and cloud.	-High false positive rate can disrupt workflows. -Complex to configure and maintain effectively.

Table 12. *Cont.*

N°	Type of Protection Mechanism	Protection Mechanism	Advantages	Disadvantages
11	Prevention	Static analysis of privacy malware [46]	-Faster and less resource-intensive than dynamic analysis. -Useful for automated scanning on endpoints and peripheral devices.	-If effective, it does not allow privacy malware to cause an initial impact. -Signature-based system. -Inefficient against zero-day attacks.
12	Prevention	IPS [103]	-If effective, it does not involve an impact.	-Fine-tuning of signatures required.
13	Response	IDS with response capabilities [67]	-Highly susceptible to false-positive cases.	-Allows threats to cause an initial impact. -Fine-tuning of signatures still required.
14	Response	Dynamic analysis of privacy malware [102]	-Allows more accurate detection and threat containment. -Offers a level of protection against zero-day attacks.	-Allows threats to cause an initial impact.
15	Recovery	Advanced Privacy Protection Systems [95]	-Able to provide details of personal data exfiltration attack trajectories and vulnerabilities exploited.	-Usually involves several subsystems or nodes. -Complex to deploy. -Usually expensive.

Figure 16 presents the number of papers by type of protection mechanism and level of assessment. While the ordinal level of assessment is widely used in detection solutions, the rest of the protection mechanisms mainly rely on the ratio level of assessment. Particularly, response and recovery solutions do not use the nominal and ordinal categories. Figure 17, on the other hand, presents the number of papers by type of protection mechanism and level of automation. For all types of protection mechanisms, automatic solutions are the majority. Notably, roughly 67% of proposed preventive solutions are automatic. However, 25% of detection solutions still rely on manual methods, mainly surveys and questionnaires, highlighting the importance of these techniques. Table 13 provides the detail of papers for each type of protection mechanism.

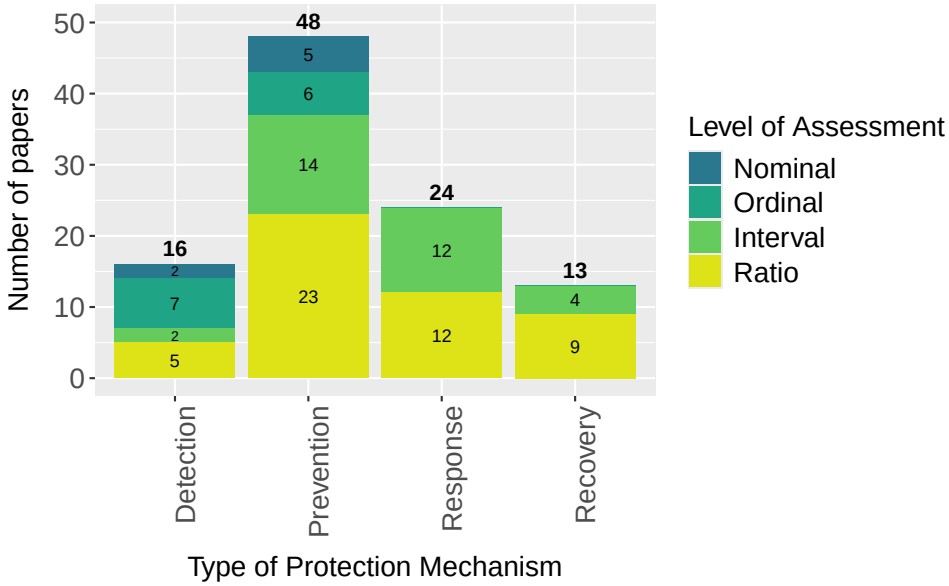

Figure 16. Number of papers by type of protection mechanism and level of assessment.

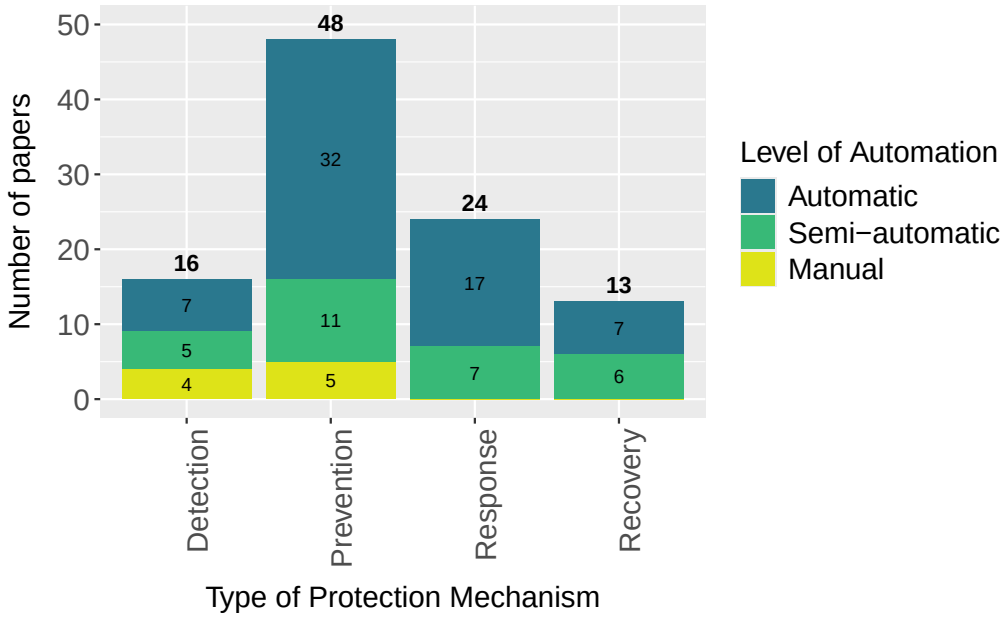

Figure 17. Number of papers by type of protection mechanism and level of automation.

Table 13. Detail of papers by type of protection mechanism.

Type of Protection Mechanism	Count	Papers
Detection	16	[1,6–8,71,84,123,137,162,168,170,172,179,181–183]
Prevention	48	[3,5,14,17,46,48,55–58,66,69,72,73,78,80,82,83,87,88,92,100,103,104,145,154,156–161,166,167,171,173–175,177,178,184,187–189,192–195]
Response	24	[16,59,63,65,67,68,70,74,79,102,146–148,151,152,155,163–165,176,185,186,190,191]
Recovery	13	[10,13,15,18,61,64,81,94,95,105,149,150,153]

5.5. Broad Domains

In this last subsection, we answer our RQ5: In what broad domains are heuristic techniques employed for the assessment of Internet privacy? Figure 15 suggests that although Internet privacy is assessed in the three broad domains of the MITRE ATT&CK [96], more research efforts are oriented toward the Enterprise broad domain, primarily driven by web technologies and the traditional use of Internet services through a web browser. Particularly, most research is focused on how to apply perception techniques, privacy policies analysis, information measurement (especially differentiation), and counting techniques in this broad domain. Significant research has also been made in the Mobile broad domain, primarily driven by mobile applications for smartphones and the IoT. However, Industrial Control Systems require more research. Figure 18 presents the number of papers by broad domain and level of assessment, level of automation, and type of protection mechanism. Notably, the number of papers for the Enterprise broad domain is higher than the number of papers for the Mobile and Industrial Control Systems broad domains in all levels of assessments, levels of automation, and types of protection mechanisms, ultimately denoting that although most heuristic techniques are based on automatic mechanisms, their applications are focused on the traditional use of Internet services through web browsers.

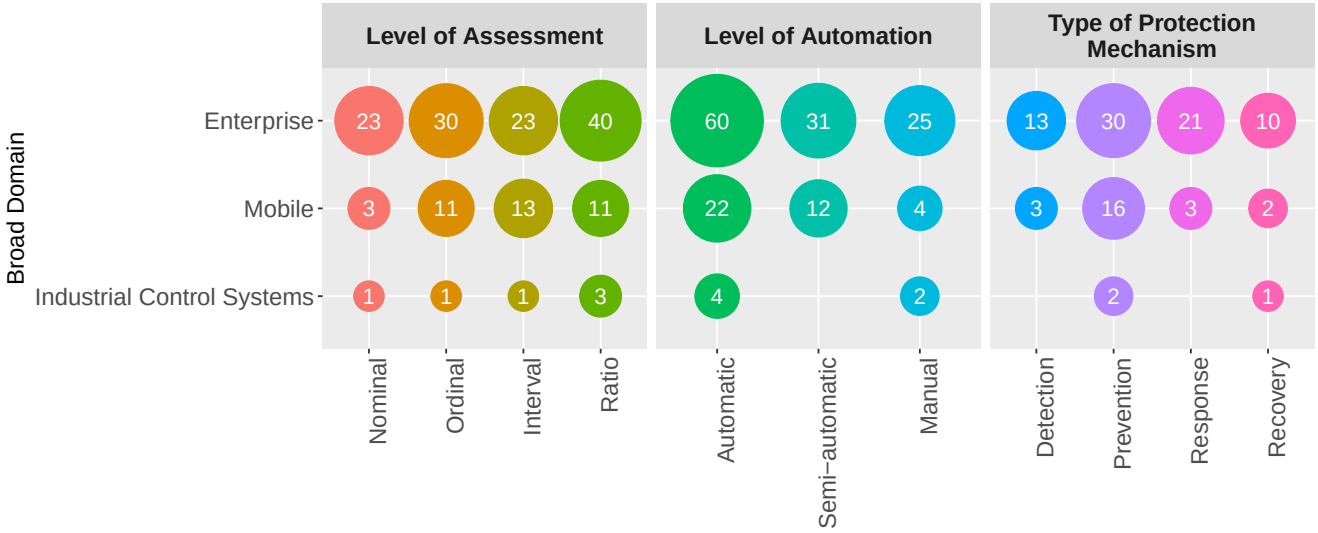

Figure 18. Number of papers by broad domain and level of assessment, level of automation, and type of protection mechanism.

Table 14 summarizes the main domains of application within each broad domain that have been widely researched for assessing Internet privacy. Table 15 offers the detail of papers categorized in each broad domain.

Table 14. Main domains of application by broad domain.

Broad Domain	Main Domains
Enterprise	Web, Cloud Computing (CC), Fog Computing (FC), Online Social Networks (OSN).
Mobile	Mobile Computing (MC), Mobile Applications (mAPPs), Internet of Things (IoT), Vehicle Ad hoc Networks (VANETs), Internet of Vehicles (IoV), Mobile Cloud Computing (MCC), Mobile Edge Computing (MEC), Mobile Health (mHealth), Mobile Payment Systems (MPS), Mobile Social Networks (MSN).
Industrial Control Systems	Industrial Internet of Things (IIoT), Industrial Automation and Control Systems (IACS), Supervisory Control and Data Acquisition (SCADA), Cyber-Physical Systems (CPS), Advanced Metering Infrastructure (AMI).

Table 15. Detail of papers by broad domain.

Broad Domain	Count	Papers
Enterprise	116	[1,3,5,7,12–15,17,18,32,47,49,50,52,56–60,62,63,65–67,69–72,74,75,78,80,81,83–87,91,94,95,100–108,110–112,114–125,127,129,132,134–142,144,146–153,155–166,170,172,173,176,179,181–192]
Mobile	38	[4,6,8–10,16,31,46,48,51,54,55,64,68,73,79,82,88,92,109,113,126,128,131,133,143,145,167–169,171,174,175,177,178,180,193,194]
Industrial Control Systems	6	[11,53,61,130,154,195]

6. Discussion

In this section, we offer an in-depth discussion of the main findings of the review, the identified gaps and challenges, and some future research directions.

6.1. Main Findings

Information measurement is one step ahead: Given the subjective nature of Internet privacy, the heuristic techniques based on measuring the amount of information disclosed by a system have been widely accepted as the most valuable and closely related to Internet

privacy [14]. The mathematical foundation on which these techniques are based is perhaps the main reason for this fact, as it provides a strong scientific basis for obtaining a valid conclusion [150]. Consequently, they have become the most extensively studied and applied techniques in the field. This trend is largely driven by the broad adoption of Differential Privacy in several studies [149]. However, other heuristic techniques, such as measuring the level of readability of privacy policies or the degree of compliance with privacy regulations they enforce, have gained significant attention, particularly due to the global enactment of privacy regulations [125]. Therefore, these alternative approaches may eventually become dominant in future research. Finally, counting techniques continue to be widely implemented, and our findings suggest they could remain valuable for assessing Internet privacy in future work [12].

Most solutions are automatic but focused on the traditional use of a web browser: The results suggest that assessing Internet privacy and implementing protection mechanisms are mostly performed by automatic solutions [18]. Although the presence of anti-bots and related controls makes the task difficult, current research demonstrates that these solutions can collect representative samples from publicly available resources, reaching significant conclusions and allowing for the generalization of results [150]. The work presented in [3] is a major example. However, they are mostly oriented toward the Enterprise broad domain, primarily due to the dominant role of web browsers in user interaction and personal data exchange [51]. Finding ways of adapting these solutions to constrained domains such as smartphones and IoT is still a promising research topic [14].

Significant research conducted manually at the nominal level of assessment: Our results indicate that meaningful research is still conducted manually to assess Internet privacy, primarily to perform a perception estimation [109]. Although nominal scales do not allow for meaningful quantitative comparisons, they remain useful for determining levels of compliance and degrees of user satisfaction [12,107]. One possible reason is the significant amount of research focused on evaluating the quality of Internet services that considers users' privacy as an evaluation criterion. The studies in [31,106,108] demonstrate why perception estimation techniques are widely adopted and are expected to remain valuable in future research.

Protecting medical data using chaotic protection mechanisms: Since the disclosure of Protected Health Information (PHI) can have severe consequences for patients, substantial research efforts have been devoted to safeguarding medical records using chaotic protection mechanisms. These mechanisms are particularly effective for Internet privacy protection, as even minimal changes in input data can result in outputs that are drastically different from previous states, ultimately enhancing unpredictability and resistance to privacy attacks. This research focus has been further reinforced by recent regulatory developments in medical data privacy, including the Health Infrastructure Security and Accountability Act (HISAA) enacted in 2024 [199] and the Healthcare Cybersecurity Act proposed in 2025 [200], both of which are grounded in the principles established by the Health Insurance Portability and Accountability Act (HIPAA) of 1996 [201].

The following are some notable examples of recent applications of chaotic systems for enhancing privacy in medical data used over the Internet. The chaotic encryption scheme proposed by Archana et al. [173] and implemented within a blockchain framework establishes secure Internet channels for the transmission of medical images over the Internet. Peng et al. [164] introduce a dual-branch framework capable of extracting only essential features from encrypted images to train chaotic neural networks. These networks leverage the inherent sensitivity and nonlinearity of chaotic dynamics, making it extremely difficult for adversaries to reconstruct the original information or infer medical patterns from the output of the chaotic neural networks. Similarly, Arévalo and Salmeron [165] propose a

privacy-preserving federated learning framework in which chaotic neural networks are trained and deployed without exposing raw medical data, enhancing the resilience of medical record protection against reverse engineering and data leakage. Finally, Kamal et al. [88] demonstrate the feasibility of implementing privacy-aware chaotic systems in the Internet of Bio-NanoThings (IoBNT) domain by introducing controlled anonymization into command signals transmitted from medical personnel to sensors implanted within the human body.

6.2. Gaps and Challenges

Lack of a standard criterion for Internet privacy assessment: Since Internet privacy is a multifaceted term difficult to define, describe, and quantify, the main challenge lies in establishing a unified criterion that clearly outlines what Internet privacy is, its components, and the scope for assessment across all existing domains. However, the subjective nature of the term, along with its exposure to evolving threats over time and the high number of privacy regulations worldwide, each with a distinct approach, makes it virtually impossible to establish a unified criterion or standard. Nevertheless, some domain-specific efforts have proposed Internet privacy assessment and compliance frameworks that can offer valuable guidance by considering the most relevant components for evaluation. For instance, Eleyan et al. [6] propose a compliance framework for assessing IoT applications communicating personal data over the Internet, limited to just four critical aspects of assessment: threat characterising, anomaly detection, dynamic policy adjustment, and regulatory compliance checking/verification. Dogra et al. [174], on the other hand, propose a framework that takes into account three aspects, including the value of user awareness, device health monitoring effectiveness, and remote management capacity. However, the study is limited to assessing Healthcare IoT applications communicating data over the Internet. Finally, the study presented in [9] proposes a framework based on the assessment of personal data transfers as a sole criterion and limited to mobile applications. This suggests that future research efforts will likely continue developing domain-specific Internet privacy assessment and compliance frameworks focused on a limited set of aspects, rather than proposing a unified criterion or standard.

Lack of defined thresholds of Internet privacy: Although research efforts have focused on applying heuristic techniques that produce measurable metrics to interpret and associate them with levels of Internet privacy, it remains unclear which threshold values constitute acceptable levels for determining whether an information system provides adequate privacy. For instance, while numerous studies, such as those exposed in [149] and [150], propose the use of Differential Privacy, there is still a lack of research establishing which ϵ values or related metrics can be considered acceptable. Similarly, machine learning classifiers used for privacy malware detection often assume that higher metrics correspond to higher levels of Internet privacy [102]. However, it remains uncertain what minimum values must be attained to effectively mitigate current threats. The evolution of threats over time complicates this scenario [2].

Privacy vs utility remains unclear: As an information system achieves higher levels of Internet privacy, the data tends to become less useful, and vice versa. There is still a paramount gap in how to appropriately associate a given level of Internet privacy with a corresponding level of data utility. Although some studies, such as the one conducted by Bousier et al. [62], have discussed this issue, they do not provide a detailed methodology for establishing this mapping. Therefore, just as heuristic techniques are mapped to specific privacy levels, further research is needed to go beyond this approach, mapping these privacy levels to corresponding levels of utility.

Misalignment between heuristic techniques and protection mechanisms: Of the 160 papers selected for the coding phase, only 101 (approximately 63%) proposed a protection mechanism that enabled the analysis and exemplification of specific heuristic techniques, for example, a Differential Privacy algorithm whose performance is assessed by measuring regret bounds between the initial and final states of an information system, as in the solution proposed by [150]; or whose effectiveness can be directly linked, using the PMA approach, to a defined level of Internet privacy, such as the privacy malware classifiers presented in [65]. Consequently, roughly 38% of the contributions are focused solely on suggesting how to assess Internet privacy, without offering any empirical observations from an implemented protection mechanism, highlighting a gap that future research must address.

The inverse situation is also evident. Many studies propose and implement protection mechanisms but fail to define evaluation criteria that link their performance to specific levels of Internet privacy, or such linkage is simply not feasible with their chosen approach. For instance, numerous ransomware machine learning classifiers are designed solely to block the malware's encryption capabilities, without addressing its personal data exfiltration behavior [163]. However, many studies that address personal data exfiltration often omit any interpretation of the achieved level of Internet privacy, focusing instead on technical aspects or merely reporting the high classification Accuracy attained [102].

Adapting statistical models for dynamic Internet privacy threats: Many heuristic techniques for Internet privacy assessment rely on the interpretation of metrics derived from statistical models and probability estimation. Prominent examples include techniques based on Differential Privacy and the Bayesian approach. However, further research is needed to determine how these models can be effectively adapted to different contexts [18]. The dynamic nature of information and the evolving threats faced on the Internet (e.g., zero-day attacks) present a significant challenge that must be addressed to achieve this goal [179]. How well these models represent real-world conditions remains a subject of ongoing debate that requires further research [158].

Lack of standardized attribute representations: Many heuristic techniques, such as those based on similarity measurement, first require defining an adequate representation of the attributes to which they are applied. Several authors have proposed a myriad of methods for this purpose through contributions such as those presented in [14,79,81], ultimately leading, for example, to the development of different similarity measurement techniques. As a result, comparing these methods becomes complex and may even be unfair. Therefore, further research is needed to develop unified approaches that enable the achievement of comparable results.

6.3. Future Research Directions

Privacy-by-default in the spotlight: While most solutions are aimed at providing a level of Internet privacy as external add-ons to an information system (e.g., web browser's plugins, algorithms executed on databases, pseudonymous communications channels, among others), creating new solutions implemented under a privacy-by-default approach will be essential, especially for compliance with privacy regulations. Therefore, research efforts will be toward achieving lightweight solutions with acceptable computational cost and that do not introduce significant latency in the provision of services over the Internet.

Telemetry modules required: Most tools that implement heuristic techniques as well as some protection mechanisms do not provide detailed information about events and correlation of events that occur in response to a privacy threat, nor do they offer continuous monitoring that allows for concluding by observing an information system over a period of time, rather than through a single observation. Even recovery solutions, such as the

adaptive mechanism proposed by Aydin and Yildirim [18], are still not able to fully operate under this approach. Furthermore, most protection mechanisms are not integrated with Security Information and Event Management (SIEM) tools, preventing their measurements from being continuously updated and included in knowledge bases for the generation of Indicators of Compromise (IoC) and Key Performance Indicators (KPI). Therefore, future research efforts will aim to propose solutions with telemetry modules to address this important issue.

Large Language Models (LLMs) to the rescue: The results of our review demonstrate that the analysis of privacy policies will continue to be widely researched. LLMs have already proven to help proposed solutions to analyze the text of these legal documents, as demonstrated by existing studies, such as the one available in [125], and work in progress, such as the research presented in [202]. Future research will likely continue leveraging this tool or its evolution, adapting its functionality to annotate privacy policies, map legal principles to regulatory frameworks, and identify privacy risks, among many other applications [203].

Future protection mechanisms: The current advancement of Artificial Intelligence (AI) and new technologies and approaches in Information Security and Privacy suggests that protection mechanisms for Internet privacy will be substantially improved in the future. Table 16 details some key technologies and approaches currently experiencing extensive research that will be paramount in protecting users' privacy when delivering services over the future Internet.

Table 16. Key technologies and approaches for enhanced protection mechanisms.

Technology/Approach	Explanation
Zero-trust networking	A security approach that assumes no user, device, or system (inside or outside the network) should be trusted by default. Access is granted based on strict identity verification, continuous monitoring, and least-privilege principles, allowing a high level of Internet privacy [204].
Intent-based privacy networking	An approach based on user-defined privacy intents (e.g., anonymity, minimal data sharing) to dynamically configure network behavior and data flows to meet those privacy goals, often powered by AI and automation [205].
AI-powered contextual integrity	A privacy framework enhanced by AI that ensures personal data is shared and used only in ways that align with the social norms, ethics, and expectations of a given context. AI systems analyze context (e.g., roles, purpose, and personal data type) to enforce appropriate personal data flows [206].
LLMs-driven privacy	Beyond privacy policies analysis, LLMs can also be used for integration into rule-based privacy systems (e.g., traditional IDS, IPS, DLP) to enhance their ability to detect, interpret, and respond to complex and context-rich threats involving personal data [207].

Privacy at the Application layer: Security and privacy schemes, such as the creation of encrypted communication channels, have commonly been applied at the Transport layer to protect data in transit [2]. These mechanisms ensure that communications between users and services are protected from interception, tampering, and eavesdropping during transmission. However, modern threats are increasingly capable of collecting personal data by executing directly on users' devices or the servers of data processors. As a result, research efforts will likely focus on proposing new protection mechanisms at the Application layer to prevent, for example, the exfiltration of sensitive information from applications that handle personal data in volatile memory.

Achieving a Defense-in-Depth (DiD) strategy would be the goal: Since Internet privacy is a complex term, relying on a single heuristic technique or a specific protection mechanism is insufficient to address the wide range of privacy threats. Therefore, the application of multiple complementary solutions layered through a DiD approach is strongly

required and likely compulsory in future research [2]. Beyond achieving excellent results through any study, it is essential to recognize that Internet privacy encompasses multiple facets, and certain aspects may be overlooked by various solutions. A comprehensive approach, integrating multiple protection mechanisms and placing the users at its core, empowered with awareness of the value of their personal data, will serve as the foundational pillar for future research [127].

7. Threats to Validity

7.1. Threats to Internal Validity

To mitigate threats to internal validity, we considered several parameters. Initially, a collaborative effort among three peers was employed for searching, screening, coding, and analyzing the identified papers. Bias was minimized through the use of pilot studies, facilitating alignment of criteria and consensus, which was evaluated using the Krippendorff's Alpha coefficient. Additionally, we developed a robust technical background to ensure a thorough understanding of each critical concept, minimizing the risk of misunderstanding during both the pilot and main phases. Lastly, to enhance clarity and consistency, we constructed a comprehensive classification scheme based on existing frameworks, complemented by a newly proposed taxonomy for classifying heuristic techniques.

7.2. Threats to External Validity

External validity was assessed through several measures. Initially, we utilized recognized databases as our primary sources of information, focusing our efforts on peer-reviewed contributions that had completed their publication stage after a rigorous review process. To ensure comprehensive coverage, we did not restrict the search to a maximum number of years, allowing us to gather all available information without omitting relevant details. On the other hand, the search terms were derived from the IEEE Thesaurus guide and meticulously analyzed before retrieving the papers from the chosen databases.

7.3. Threats to Conclusion

The measures implemented to address threats to both internal and external validity served as foundational pillars that validated the research conclusion. This validity is substantiated by the systematic and staged methodology employed, as well as by anchoring our approach in existing contributions that comprehensively review the literature on the assessment of Internet privacy and related protection mechanisms.

8. Conclusions

This paper has comprehensively reviewed heuristic techniques for assessing Internet privacy and related protection mechanisms. Overall, most research focuses on applying heuristic techniques based on information measurement, primarily driven by the use of Differential Privacy. However, the analysis of privacy policies is experiencing significant growth, and counting techniques are expected to remain widely used in the future. Although most heuristic techniques are based on automatic mechanisms, they are largely applied within the traditional context of Internet services accessed through a web browser. This highlights the need for further research in other domains.

As future work, efforts will be directed toward proposing privacy technologies based on a privacy-by-default approach to ensure compliance with privacy regulations. In addition, these technologies will include telemetry modules capable of integrating with SIEM solutions and suggesting context-aware metrics derived from observations over time, going beyond single-point assessments. The implementation of strategies based on a DiD

approach will therefore be crucial to safeguarding the privacy of users who access Internet services in the face of evolving threats over time.

Author Contributions: Conceptualization, D.C.-S.; methodology, D.C.-S. and D.S.G.; software, D.C.-S. and D.S.G.; validation, J.E.-J. and D.S.G.; formal analysis, D.C.-S.; investigation, D.C.-S.; resources, J.E.-J.; data curation, D.C.-S., J.E.-J. and D.S.G.; writing—original draft preparation, D.C.-S.; writing—review and editing, D.C.-S.; visualization, D.C.-S.; supervision, J.E.-J. and D.S.G.; project administration, J.E.-J.; funding acquisition, J.E.-J. All authors have read and agreed to the published version of the manuscript.

Funding: The publication of the results of this research is funded by the Escuela Politécnica Nacional, in particular through the project "PRIVIA: Identificación Automatizada de Brechas de Privacidad en Ecuador usando Inteligencia Artificial Generativa y LLMs", ref. PIGR-24-06.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the files required for reproducibility of experiments are publicly available at https://github.com/dcevallossalas/InternetPrivacyAssessment (accessed on 23 June 2025).

Acknowledgments: The authors gratefully acknowledge the financial support provided by the Escuela Politécnica Nacional.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
APIs	Application Programming Interfaces
CC	Cloud Computing
CPS	Cyber-Physical Systems
CPPA	Consumer Privacy Protection Act
CPRA	California Privacy Rights Act
DiD	Defense-in-Depth
DLP	Data Loss Prevention
FC	Fog Computing
GDPR	General Data Protection Regulation
HIPAA	Health Insurance Portability and Accountability Act
HISAA	Health Infrastructure Security and Accountability Act
HTTP	Hypertext Transfer Protocol
IACS	Industrial Automation and Control Systems
IDS	Intrusion Detection Systems
IoBNT	Internet of Bio-NanoThings
IIoT	Industrial Internet of Things
IoC	Indicators of Compromise
IoT	Internet of Things
IoUT	Internet of Underwater Things
IoV	Internet of Vehicles
IPCs	Internet Privacy Concerns
IPS	Intrusion Prevention Systems
KPI	Key Performance Indicators
LLMs	Large Language Models
MAE	Mean Absolute Error

MC	Mobile Computing
MCC	Mobile Cloud Computing
MEC	Mobile Edge Computing
mAPPs	Mobile Applications
mGovernment	Mobile Government
mHealth	Mobile Health
MPS	Mobile Payment Systems
MSN	Mobile Social Networks
MSE	Mean Squared Error
OSN	Online Social Networks
OT	Optimal Transport
PDA	Personal Digital Assistants
PETs	Privacy-enhancing technologies
PHI	Protected Health Information
PMA	Protection Mechanism Approach
RMSE	Root Mean Squared Error
RQs	Research Questions
SCADA	Supervisory Control and Data Acquisition
SIEM	Security Information and Event Management
VANETs	Vehicle Ad hoc Networks

References

1. Jemili , F.; Jouini, K.; Korbaa, O. Intrusion Detection based on Concept Drift Detection & Online Incremental Learning. *Int. J. Pervasive Comput. Commun.* **2025**, *21*, 81–115. [CrossRef]
2. Cevallos-Salas, D.; Estrada-Jiménez, J.; Guamán, D.S. Application layer security for Internet communications: A comprehensive review, challenges, and future trends. *Comput. Electr. Eng.* **2024**, *119*, 109498. [CrossRef]
3. Pansy, D.; Mahalakshmi, S.; Vidya, B. End-To-End Encryption Solution for Social Media Networks. In Proceedings of the International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI 2025), Chennai, India, 28–29 March 2025; pp. 1–6. [CrossRef]
4. Zafar, A.; Das, A. Comparative Privacy Analysis of Mobile Browsers. In Proceedings of the 13th ACM Conference on Data and Application Security and Privacy (CODASPY 2023), Charlotte, NC, USA, 24–26 April 2023; pp. 3–14. [CrossRef]
5. Ormeno, M.; Dao, H.; Herskovic, V.; Fukuda, K. Do Cookie Banners Respect My Browsing Privacy? Measuring the Effectiveness of Cookie Rejection for Limiting Behavioral Advertising. *IEEE Access* **2024**, *12*, 174539–174550. [CrossRef]
6. Lung, L.W.; Wang, Y.R.; Chen, Y.S. Leveraging Deep Learning and Internet of Things for Dynamic Construction Site Risk Management. *Buildings* **2025**, *15*, 1325. [CrossRef]
7. Abdullah Punjabi, T.; Alban, A.Q.; Barhamgi, M. Effective Transparent Monitoring of Personal Data. In Proceedings of the Cyberspace Simulation and Evaluation (CSE 2024), Doha, Qatar, 2–5 December 2024; Springer: Singapore, 2025; pp. 29–45. [CrossRef]
8. Bemmann, F.; Stoll, H.; Mayer, S. Privacy Slider: Fine-Grain Privacy Control for Smartphones. *Proc. ACM Hum.-Comput. Interact.* **2024**, *8*, 1–31. [CrossRef]
9. Rodriguez, D.; Del Alamo, J.M.; Fernández-Aller, C.; Sadeh, N. Sharing is Not Always Caring: Delving Into Personal Data Transfer Compliance in Android Apps. *IEEE Access* **2024**, *12*, 5256–5269. [CrossRef]
10. Singla, P.; Garg, H.; Gagandeep; Pathak, A.; Singh, S.P. Privacy Enhancement in Internet of Things (IoT) via mRMR for prevention and avoidance of data leakage. *Comput. Electr. Eng.* **2024**, *116*, 109151. [CrossRef]
11. Ceci, J.; Stegman, J.; Khan, H. No Privacy in the Electronics Repair Industry. In Proceedings of the IEEE Symposium on Security and Privacy (IEEE S&P 2023), San Francisco, CA, USA, 21–25 May 2023; Volume 2023, pp. 3347–3364. [CrossRef]
12. Paguay-Chimarro, C.; Cevallos-Salas, D.; Rodríguez-Hoyos, A.; Estrada-Jiménez, J. Transparency Unleashed: Privacy Risks in the Age of E-Government. *Informatics* **2025**, *12*, 39. [CrossRef]
13. Luceri, L.; Andreoletti, D.; Tornatore, M.; Braun, T.; Giordano, S. Measurement and control of geo-location privacy on Twitter. *Online Soc. Netw. Media* **2020**, *17*, 100078. [CrossRef]
14. Kara, B.C.; Eyupoglu, C.; Karakuş, O. (r, k, e)-Anonymization: Privacy-Preserving Data Publishing Algorithm Based on Multi-Dimensional Outlier Detection, k-Anonymity, and e-Differential Privacy. *IEEE Access* **2025**, *13*, 70422–70435. [CrossRef]
15. Xiao, H.; Devadas, S. PAC Privacy: Automatic Privacy Measurement and Control of Data Processing. In *Proceedings of the Advances in Cryptology—CRYPTO 2023. CRYPTO 2023*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2023; Volume 14082, pp. 611–644. [CrossRef]

16. Alhalabi, W.; Al-Rasheed, A.; Manoharan, H.; Alabdulkareem, E.; Alduailij, M.; Alduailij, M.; Selvarajan, S. Distinctive Measurement Scheme for Security and Privacy in Internet of Things Applications Using Machine Learning Algorithms. *Electronics* **2023**, *12*, 747. [CrossRef]
17. Anderson, E.; Chase, M.; Durak, F.B.; Laine, K.; Weng, C. Precio: Private Aggregate Measurement via Oblivious Shuffling. In Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS 2024), CCS '24, Salt Lake City, UT, USA, 14–18 October 2024; pp. 1819–1833. [CrossRef]
18. Aydin, S.; Yıldırım, S. Bayesian Frequency Estimation under Local Differential Privacy with an Adaptive Randomized Response Mechanism. *ACM Trans. Knowl. Discov. Data* **2025**, *19*, 1–40. [CrossRef]
19. Kovacs, B. Data Breach at Healthcare Services Firm Episource Impacts 5.4 Million People. Available online: https://www.securityweek.com/data-breach-at-healthcare-services-firm-episource-impacts-5-4-million-people (accessed on 22 July 2025).
20. Hurley, J. Over 5.4 Million US Patient Records Compromised in Episource Data Breach. Available online: https://www.grip.globalrelay.com/over-5-4-million-us-patient-records-compromised-in-episource-data-breach (accessed on 22 July 2025).
21. McAfee. RockYou2024: Unpacking the Largest Password Leak in History. Available online: https://www.mcafee.com/blogs/internet-security/rockyou2024-unpacking-the-largest-password-leak-in-history (accessed on 22 July 2025).
22. America's Cyber Defense Agency. #StopRansomware: CL0P Ransomware Gang Exploits CVE-2023-34362 MOVEit Vulnerability. Available online: https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-158a (accessed on 22 July 2025).
23. Datta, P.M.; Acton, T. Ransomware and Costa Rica's national emergency: A defense framework and teaching case. *J. Inf. Technol. Teach. Cases* **2024**, *14*, 56–67. [CrossRef]
24. Yun, H.; Lee, G.; Kim, D.J. A meta-analytic review of empirical research on online information privacy concerns: Antecedents, outcomes, and moderators. In Proceedings of the 35th International Conference on Information Systems "Building a Better World Through Information Systems (ICIS 2014)", Auckland, New Zealand, 14–17 December 2014; Volume 2014.
25. Bartol, J.; Vehovar, V.; Petrovčič, A. Should we be concerned about how information privacy concerns are measured in online contexts? A systematic review of survey scale development studies. *Informatics* **2021**, *8*, 31. [CrossRef]
26. Bartol, J.; Vehovar, V.; Petrovčič, A. Systematic review of survey scales measuring information privacy concerns on social network sites. *Telemat. Inform.* **2023**, *85*, 102063. [CrossRef]
27. van der Schyff, K.; Prior, S.; Renaud, K. Privacy policy analysis: A scoping review and research agenda. *Comput. Secur.* **2024**, *146*, 104065. [CrossRef]
28. Del Alamo, J.; Guaman, D.; García, B.; Díez Medialdea, A. A systematic mapping study on automated analysis of privacy policies. *Computing* **2022**, *104*, 2053–2076. [CrossRef]
29. Majeed, A.; Khan, S.; Hwang, S.O. Toward Privacy Preservation Using Clustering Based Anonymization: Recent Advances and Future Research Outlook. *IEEE Access* **2022**, *10*, 53066–53097. [CrossRef]
30. Barajas, J.; Bhamidipati, N.; Shanahan, J.G. Online Advertising Incrementality Testing: Practical Lessons, Paid Search and Emerging Challenges. In *Advances in Information Retrieval. ECIR 2022*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2022; Volume 13186, pp. 575–581. [CrossRef]
31. Alghareeb, M.; Albesher, A.S.; Asif, A. Studying Users' Perceptions of COVID-19 Mobile Applications in Saudi Arabia. *Sustainability* **2023**, *15*, 956. [CrossRef]
32. Swartz, P.; Da Veiga, A.; Martins, N. Validating an information privacy governance questionnaire to measure the perception of employees. *Inf. Comput. Secur.* **2021**, *29*, 761–786. [CrossRef]
33. Esposito, C.; Moscato, V.; Sperlì, G. Detecting malicious reviews and users affecting social reviewing systems: A survey. *Comput. Secur.* **2023**, *133*, 103407. [CrossRef]
34. Eke, C.I.; Norman, A.A.; Shuib, L.; Nweke, H.F. A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions. *IEEE Access* **2019**, *7*, 144907–144924. [CrossRef]
35. Senette, C.; Siino, M.; Tesconi, M. User Identity Linkage on Social Networks: A Review of Modern Techniques and Applications. *IEEE Access* **2024**, *12*, 171241–171268. [CrossRef]
36. Desmal, A.J.; Hamid, S.; Othman, M.K.; Zolait, A. A user satisfaction model for mobile government services: A literature review. *PeerJ Comput. Sci.* **2022**, *8*, e1074. [CrossRef]
37. Youm, H.Y. An overview of security and privacy issues for internet of thing. *IEICE Trans. Inf. Syst.* **2017**, *E100D*, 1649–1662. [CrossRef]
38. Haque, M.A.; Haque, S.; Kumar, K.; Rahman, M.; Sonal, D.; Almrezeq, N. Security and Privacy in Internet of Things. *Commun. Comput. Inf. Sci.* **2022**, *1591 CCIS*, 182–196. [CrossRef]
39. Dong, R.; Ratliff, L.J.; Cárdenas, A.A.; Ohlsson, H.; Sastry, S.S. Quantifying the utility–privacy tradeoff in the internet of things. *ACM Trans. Cyber-Phys. Syst.* **2018**, *2*, 1–28. [CrossRef]
40. Bhattacharjya, K.; De, D. Federated learning-based privacy-preserving Internet of Underwater Things: A vision, architecture, computing, taxonomy, and future directions. *J. Supercomput.* **2025**, *81*, 870. [CrossRef]

41. Benamor, Z.; Seghir, Z.A.; Djezzar, M.; Hemam, M. A comparative study of machine learning algorithms for intrusion detection in IoT networks. *Rev. D'Intell. Artif.* **2023**, *37*, 567–576. [CrossRef]
42. Al-Qarni, E.A. Cybersecurity in Healthcare: A Review of Recent Attacks and Mitigation Strategies. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 135–140. [CrossRef]
43. Hameed, S.S.; Hassan, W.H.; Latiff, L.A.; Ghabban, F. A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches. *PeerJ Comput. Sci.* **2021**, *7*, e414. [CrossRef]
44. Paul, S.K.; Knox, D. A Taxonomy and Gap-Analysis in Digital Privacy Education. In *Foundations and Practice of Security. FPS 2022*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2023; Volume 13877, pp. 221–235. [CrossRef]
45. Rajkumar, K.; Dhanakoti, V. Methodological Survey to Improve the Secure Data Storage in Cloud Computing. In Proceedings of the International Conference on Emerging Smart Computing and Informatics (ESCI 2020), Pune, India, 12–14 March 2020; Volume 2020, pp. 313–317. [CrossRef]
46. Chhetri, C.; Motti, V.G. Enhancing the design of data-related privacy controls for smart home devices. *Inf. Comput. Secur.* **2023**, *31*, 368–384. [CrossRef]
47. Ghaith, M.M.; Alzuod, M.A.; Al-Okaily, M. The Effect of Information Sensitivity and Usage on Consumer Privacy: The Case of Electronic Payment Systems. In *From the Internet of Things to the Internet of Ideas: The Role of Artificial Intelligence. EAMMIS 2022*; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2023; Volume 557, pp. 391–401. [CrossRef]
48. Gadiyar, H.M.T.; Thyagaraju, G.; Goudar, R. An Adaptive Approach for Preserving Privacy in Context Aware Applications for Smartphones in Cloud Computing Platform. *Int. J. Adv. Comput. Sci. Appl.* **2022**, *13*, 519–529. [CrossRef]
49. Ding, H.; Zhang, S.; Zhou, L.; Yang, P. Readability Analysis of Privacy Policies for Large-Scale Websites: A Perspective from Deep Learning and Linguistics. In Proceedings of the IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta 2022), Haikou, China, 15–18 December 2022; pp. 1748–1753. [CrossRef]
50. Esmaeilzadeh, P. The effect of the privacy policy of Health Information Exchange (HIE) on patients' information disclosure intention. *Comput. Secur.* **2020**, *95*, 101819. [CrossRef]
51. Guamán, D.S.; Rodriguez, D.; del Alamo, J.M.; Such, J. Automated GDPR compliance assessment for cross-border personal data transfers in android applications. *Comput. Secur.* **2023**, *130*, 103262. [CrossRef]
52. Avorgbedor, F.; Liu, J. Enhancing User Privacy Protection by Enforcing Clark-Wilson Security Model on Facebook. In Proceedings of the IEEE International Conference on Electro Information Technology (EIT 2020), Chicago, IL, USA, 31 July–1 August 2020; Volume 2020, pp. 155–161. [CrossRef]
53. Li, Z.; Oechtering, T.J.; Gunduz, D. Privacy against a hypothesis testing adversary. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 1567–1581. [CrossRef]
54. Bi, R.; Chen, Q.; Chen, L.; Xiong, J.; Wu, D. A Privacy-Preserving Personalized Service Framework through Bayesian Game in Social IoT. *Wirel. Commun. Mob. Comput.* **2020**, *2020*, 8891889. [CrossRef]
55. Wu, Y.; Peng, H.; Zhang, X.; Chen, H.; Li, C. Publish me and protect me: Personalized and flexible location privacy protection in mobile social networks. In Proceedings of the IEEE 23rd International Symposium on Quality of Service (IWQoS 2016), Beijing, China, 20–21 June 2016; Volume 1, pp. 147–152. [CrossRef]
56. Motahari, S.; Ziavras, S.G.; Jones, Q. Online anonymity protection in computer-mediated communication. *IEEE Trans. Inf. Forensics Secur.* **2010**, *5*, 570–580. [CrossRef]
57. Martínez Rodríguez, D.; Nin, J.; Nuñez-del Prado, M. Towards the adaptation of SDC methods to stream mining. *Comput. Secur.* **2017**, *70*, 702–722. [CrossRef]
58. Zhang, W.; Li, M.; Tandon, R.; Li, H. Online location trace privacy: An information theoretic approach. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 235–250. [CrossRef]
59. Li, W.; Li, H. LRDM: Local record-driving mechanism for big data privacy preservation in social networks. In Proceedings of the IEEE 1st International Conference on Data Science in Cyberspace (DSC 2017), Shenzhen, China, 26–29 June 2017; Volume 1, pp. 556–560. [CrossRef]
60. Jain, P.; Kothari, P.; Thakurta, A. Differentially private online learning. *J. Mach. Learn. Res.* **2012**, *23*, 24.1–24.34.
61. Yang, L.; Chen, X.; Zhang, J.; Poor, H.V. Optimal privacy-preserving energy management for smart meters. In Proceedings of the IEEE INFOCOM (2014), Toronto, ON, Canada, 27 April–2 May 2014; Volume 1, pp. 513–521. [CrossRef]
62. Boursier, E.; Perchet, V. Utility/Privacy Trade-off through the lens of Optimal Transport. In Proceedings of the Machine Learning Research (PMLR 2020), Virtual, 13–18 July 2020; Volume 108, pp. 591–601.
63. Cortese, A.; Masoumzadeh, A. Modeling exposure in online social networks. In Proceedings of the 15th Annual Conference on Privacy, Security and Trust (PST 2018), Belfast, UK, 28–30 August 2018; Volume 2018, pp. 327–336. [CrossRef]
64. Zhang, H.; Xu, Z.; Zhou, Z.; Shi, J.; Du, X. CLPP: Context-aware location privacy protection for location-based social network. In Proceedings of the IEEE International Conference on Communications (ICC 2015), London, UK, 8–12 June 2015; Volume 2015, pp. 1164–1169. [CrossRef]

65. Sirigiri, M.; Sirigiri, D.; Aishwarya, R.; Yogitha, R. Malware Detection and Analysis using Machine Learning. In Proceedings of the 7th International Conference on Computing Methodologies and Communication (ICCMC 2023), Erode, India, 23–25 February 2023; Volume 1, pp. 1074–1081. [CrossRef]
66. Chandrasekhar, T.; Kumar, S. A Noval Method for Cloud Security and Privacy Using Homomorphic Encryption Based on Facial Key Templates. *J. Adv. Inf. Technol.* **2022**, *13*, 638–644. [CrossRef]
67. Al-Ghuwairi, A.R.; Sharrab, Y.; Al-Fraihat, D.; AlElaimat, M.; Alsarhan, A.; Algarni, A. Intrusion detection in cloud computing based on time series anomalies utilizing machine learning. *J. Cloud Comput.* **2023**, *12*, 127. [CrossRef]
68. Ajjaj, S.; El Houssaini, S.; Hain, M.; El Houssaini, M.A. Incremental Online Machine Learning for Detecting Malicious Nodes in Vehicular Communications Using Real-Time Monitoring. *Telecom* **2023**, *4*, 629–648. [CrossRef]
69. Sanchez-Rola, I.; Balzarotti, D.; Santos, I. Cookies from the Past: Timing Server-side Request Processing Code for History Sniffing. *Digit. Threat. Res. Pract.* **2020**, *1*, 1–24. [CrossRef]
70. Sanchez-Rola, I.; Balzarotti, D.; Santos, I. BakingTimer: Privacy analysis of server-side request processing time. In Proceedings of the ACM International Conference Proceeding Series (ICPS 2019), San Juan, PR, USA, 9–13 December 2019; Volume 1, pp. 478–488. [CrossRef]
71. Solomos, K.; Ilia, P.; Ioannidis, S.; Kourtellis, N. Talon: An automated framework for cross-device tracking detection. In Proceedings of the 22nd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2019), Beijing, China, 23–25 September 2019; pp. 227–241.
72. Zhong, K.; Ma, Y.; Angel, S. Ibex: Privacy-preserving Ad Conversion Tracking and Bidding. In Proceedings of the ACM Conference on Computer and Communications Security (ACM CCS 2022), Los Angeles, CA, USA, 7–11 November 2022; Volume 1, pp. 3223–3237. [CrossRef]
73. Machiko Mada, M.; Razak Saputri, N.; Ayu Br Keliat, S.; Nugroho Perwiro Atmojo, R. The Analysis of Advertising Identification Through Measurement of User Activity on Smartphone Devices. In Proceedings of the 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability (ICITISEE 2022), Yogyakarta, Indonesia, 13–14 December 2022; Volume 1, pp. 437–442. [CrossRef]
74. Yokoyama, S.; Uda, R. A proposal of preventive measure of pursuit using a browser fingerprint. In Proceedings of the ACM Ubiquitous Information Management and Communication (IMCOM 2015), Bali, Indonesia, 8–10 January 2015; Volume 1, pp. 1–7. [CrossRef]
75. Shariatnasab, M.; Shirani, F.; Erkip, E. Fundamental Privacy Limits in Bipartite Networks Under Active Attacks. *IEEE J. Sel. Areas Commun.* **2022**, *40*, 940–954. [CrossRef]
76. for Standardization, I.O. Information Security, Cybersecurity and Privacy Protection—Privacy Information Management Systems—Requirements and Guidance. Available online: https://www.iso.org/standard/85819.html (accessed on 1 July 2025).
77. National Institute of Standards and Technology. NIST Privacy Framework: A Tool for Improving Privacy Through Enterprise Risk Management, Version 1.0. Available online: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf (accessed on 29 May 2025).
78. Shaik, A.; Masilamani, V. Zero watermarking scheme for privacy protection in e-Health care. *Automatika* **2023**, *64*, 453–466. [CrossRef]
79. Kilany, R.; Sokhn, M.; Hellani, H.; Shabani, S. Towards flexible K-anonymity. *Commun. Comput. Inf. Sci.* **2016**, *649*, 288–297. [CrossRef]
80. Arampatzis, A.; Drosatos, G.; Efraimidis, P.S. Versatile Query Scrambling for Private Web Search. *Inf. Retr.* **2015**, *18*, 331–358. [CrossRef]
81. Kaur, H.; Hooda, N.; Singh, H. k-anonymization of social network data using Neural Network and SVM: K-NeuroSVM. *J. Inf. Secur. Appl.* **2023**, *72*, 103382. [CrossRef]
82. Shin, M.; Cornelius, C.; Kapadia, A.; Triandopoulos, N.; Kotz, D. Location privacy for mobile crowd sensing through population mapping. *Sensors* **2015**, *15*, 15285–15310. [CrossRef]
83. Lavesson, N.; Johnson, H. Measuring profile distance in online social networks. In Proceedings of the ACM International Conference Proceeding Series (ICPS 2011), Sogndal, Norway, 25–27 May 2011; Volume 1, pp. 1–12. [CrossRef]
84. Kou, H.; Wang, F.; Lv, C.; Dong, Z.; Huang, W.; Wang, H.; Liu, Y. Trust-Based Missing Link Prediction in Signed Social Networks with Privacy Preservation. *Wirel. Commun. Mob. Comput.* **2020**, *2020*, 1–10. [CrossRef]
85. Kabwe, F.; Phiri, J. Identity attributes metric modelling based on mathematical distance metrics models. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 450–464. [CrossRef]
86. Coscia, M.; Mendez-Bermudez, A. Pearson correlations on complex networks. *J. Complex Netw.* **2021**, *9*, 1–14. [CrossRef]
87. Silva, P.; Kencl, L.; Monteiro, E. Data Privacy Protection—Concealing Text and Audio with a DNA-inspired Algorithm. In Proceedings of the 12th International Conference on Autonomous Infrastructure, Management and Security (AIMS 2018), Munich, Germany, 4–5 June 2018; Volume 1, pp. 46–59.

88. Kamal, I.R.; Abd El-atty, S.; El-Zoghdy, S.; Soliman, R. Internet of Bio-NanoThings privacy: Securing a multi compartmental targeted cancer drug delivery scheme. *Multimed. Tools Appl.* **2024**, *83*, 79235–79258. [CrossRef]
89. Stevens, S.S. On the Theory of Scales of Measurement. *Science* **1946**, *103*, 677–680. [CrossRef] [PubMed]
90. National Institute of Standards and Technology. Measurement Guide for Information Security: Volume 1—Identifying and Selecting Measures. Available online: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-55v1.pdf (accessed on 25 June 2025).
91. Romano, R.; Han, J. Consumer Perceptions towards Unsolicited Advertisements on Social Media. *Data* **2022**, *7*, 138. [CrossRef]
92. Thota, C.; Mavromoustakis, C.; Mastorakis, G. CAP2M.÷Contingent Anonymity Preserving Privacy Method for the Internet of Things Services. *Comput. Electr. Eng.* **2023**, *107*, 108640. [CrossRef]
93. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. Available online: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf (accessed on 26 June 2025).
94. Mai, W.; Xiao, Y. A novel GPU based Geo-Location Inference Attack on WebGL framework. *High-Confid. Comput.* **2023**, *3*, 100135. [CrossRef]
95. Satish Kumar, A.; Revathy, S. A hybrid soft computing with big data analytics based protection and recovery strategy for security enhancement in large scale real world online social networks. *Theor. Comput. Sci.* **2022**, *927*, 15–30. [CrossRef]
96. MITRE Corporation. MITRE ATT&CK®. Available online: https://attack.mitre.org/ (accessed on 29 May 2025).
97. IEEE. Access the IEEE thesaurus. Available online: https://www.ieee.org/publications/services/thesaurus-access-page.html (accessed on 25 June 2025).
98. The California Privacy Rights Act of 2020. The CPRA Organization. Available online: https://thecpra.org/ (accessed on 5 July 2025).
99. Government of Canada. Consumer Privacy Protection Act. Available online: https://ised-isde.canada.ca/site/innovation-better-canada/en/consumer-privacy-protection-act (accessed on 5 July 2025).
100. Yang, D.; Qu, B.; Cudré-Mauroux, P. Privacy-Preserving Social Media Data Publishing for Personalized Ranking-Based Recommendation. *IEEE Trans. Knowl. Data Eng.* **2019**, *31*, 507–520. [CrossRef]
101. Zhang, J.; Shi, S.T.; Weng, C.J.; Xu, L. Individual Attribute and Cascade Influence Capability-Based Privacy Protection Method in Social Networks. *Secur. Commun. Netw.* **2022**, *2022*, 6338123. [CrossRef]
102. Cevallos-Salas, D.; Grijalva, F.; Estrada-Jiménez, J.; Benítez, D.; Andrade, R. Obfuscated Privacy Malware Classifiers Based on Memory Dumping Analysis. *IEEE Access* **2024**, *12*, 17481–17498. [CrossRef]
103. Al-Sarayrah, T.A.; Li, Z.; Yin, L.; Mostafa, A.M. Online Verification and Enforcement of Sequential K-Opacity in Extended Probabilistic Automata. *IEEE Access* **2024**, *12*, 84189–84203. [CrossRef]
104. Hamzehzadeh, S.; Mazinani, S.M. ANNM: A New Method for Adding Noise Nodes Which are Used Recently in Anonymization Methods in Social Networks. *Wirel. Pers. Commun.* **2019**, *107*, 1995–2017. [CrossRef]
105. Gao, W.; Liang, W.; Hao, R.; Yu, J. Enabling privacy-preserving non-interactive computation for Hamming distance. *Inf. Sci.* **2024**, *670*, 120592. [CrossRef]
106. Muhammad, M.Z.; Muhamad, F.H.; Doktoralina, C.M.; Mukhtar, D.; Ghazali, M.F.; Rahman, M.K.; Din, N.M. Online Banking of Islamic Banks: The Customer's Perceptions in Kelantan. In *Sustainable Finance, Digitalization and the Role of Technology. ICBT 2021*; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2023; Volume 487, pp. 859–873. [CrossRef]
107. Chakraborty, R.; Prakasha, G.; Sripavithra, C. Factors Affecting Data-Privacy Protection and Promotion of Safe Digital Usage. In Proceedings of the CEUR Workshop Proceedings (CEUR-WS.org 2022), Ternopil, Ukraine, 22–24 November 2022; Volume 3094, pp. 49–58.
108. Gabrielli, S.; Rizzi, S.; Mayora, O.; More, S.; Baun, J.C.P.; Vandevelde, W. Multidimensional Study on Users' Evaluation of the KRAKEN Personal Data Sharing Platform. *Appl. Sci.* **2022**, *12*, 3270. [CrossRef]
109. Joy, D.; Kotevska, O.; Al-Masri, E. Investigating Users' Privacy Concerns of Internet of Things (IoT) Smart Devices. In Proceedings of the IEEE Eurasia Conference on IoT, Communication and Engineering (ECICE 2022), Yunlin, Taiwan, 28–30 October 2022; Volume 1, pp. 70–76. [CrossRef]
110. Liu, Y.; Tse, W.K.; Kwok, P.Y.; Chiu, Y.H. Impact of Social Media Behavior on Privacy Information Security Based on Analytic Hierarchy Process. *Information* **2022**, *13*, 280. [CrossRef]
111. Aji, Z.M.; Mohd Salleh, N.S.; Zakaria, N.H.; Mohd Khalid, A.H. Are You Aware of Your Data Privacy? The Case of Digital Contact Tracing Applications (MySejahtera). In Proceedings of the 2021 7th International Conference on Research and Innovation in Information Systems (ICRIIS 2021), Johor Bahru, Malaysia, 25–26 October 2021; Volume 1, pp. 1–6. [CrossRef]
112. McGill, T.; Thompson, N. Exploring potential gender differences in information security and privacy. *Inf. Comput. Secur.* **2021**, *29*, 850–865. [CrossRef]
113. Sharma, T.; Dyer, H.A.; Bashir, M. Enabling User-centered Privacy Controls for Mobile Applications: COVID-19 Perspective. *ACM Trans. Internet Technol.* **2021**, *21*, 1–24. [CrossRef]

114. AlMarzooqi, F.M.; Moonesar, I.A.; AlQutob, R. Healthcare professional and user perceptions of eHealth data and record privacy in Dubai. *Information* **2020**, *11*, 415. [CrossRef]
115. Son, Y.; Cho, G.; Kim, H.; Woo, S. Understanding users' risk perceptions about personal health records shared on social networking services. In Proceedings of the ACM Asia Conference on Computer and Communications Security (AsiaCCS 2019), Auckland, New Zealand, 9–12 July 2019; Volume 1, pp. 352–365. [CrossRef]
116. Suen, H.Y. The effects of employer SNS motioning on employee perceived privacy violation, procedural justice, and leave intention. *Ind. Manag. Data Syst.* **2018**, *118*, 1153–1169. [CrossRef]
117. Bauer, L.; Bravo-Lillo, C.; Fragkaki, E.; Melicher, W. A comparison of users' perceptions of and willingness to use Google, Facebook, and Google+ single-sign-on functionality. In Proceedings of the ACM Conference on Computer and Communications Security (CCS 2013), Berlin, Germany, 8 November 2013; Volume 1, pp. 25–36. [CrossRef]
118. Carrascal, J.P.; Riederer, C.; Erramilli, V.; Cherubini, M.; De Oliveira, R. Your browsing behavior for a big mac: Economics of personal information online. In Proceedings of the 22nd International Conference on World Wide Web (WWW 2013), Rio de Janeiro, Brazil, 13–17 May 2013; Volume 1, pp. 189–199.
119. Bergmann, M. Testing privacy awareness. *IFIP Adv. Inf. Commun. Technol.* **2009**, *298*, 237–253. [CrossRef]
120. Joinson, A.N.; Paine, C.; Buchanan, T.; Reips, U.D. Measuring self-disclosure online: Blurring and non-response to sensitive items in web-based surveys. *Comput. Hum. Behav.* **2008**, *24*, 2158–2171. [CrossRef]
121. Ragab, A.; Mannan, M.; Youssef, A. "Trust Me Over My Privacy Policy": Privacy Discrepancies in Romantic AI Chatbot Apps. In Proceedings of the 2024 IEEE European Symposium on Security and Privacy Workshops (EuroSPW 2024), Vienna, Austria, 8–12 July 2024; pp. 484–495. [CrossRef]
122. Zac, A.; Wey, P.; Bechtold, S.; Rodriguez, D.; Del Alamo, J. The Court Speaks, But Who Listens? Automated Compliance Review of the GDPR. *SSRN Electron. J.* **2024** . [CrossRef]
123. Alghoul, M.; Moutan, M.; Kastro, N.; Karakra, A.; Barghouthi, H.; Njoum, M.; Awadallah, R. Policy Peek: Privacy Policy Automatic Analysis Tool. In Proceedings of the International Conference on Electrical, Computer and Energy Technologies (ICECET 2024), Sydney, Australia, 25–27 July 2024; pp. 1–6. [CrossRef]
124. Maraba, J.; Veiga, A. A Study of Online Privacy Policies of South African Retail Websites. In Proceedings of the International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability (ARTIIS 2024), Santiago de Chile, Chile, 21–23 October 2024; pp. 426–440. [CrossRef]
125. Rodriguez, D.; Yang, I.; Del Alamo, J.; Sadeh, N. Large language models: A new approach for privacy policy analysis at scale. *Computing* **2024**, *106*, 3879–3903. [CrossRef]
126. Rodríguez, D.; Fernández-Aller, C.; Del Alamo, J.M.; Sadeh, N. Data Retention Disclosures in the Google Play Store: Opacity Remains the Norm. In Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroSPW 2024), Vienna, Austria, 8–12 July 2024; pp. 19–23. [CrossRef]
127. Vu, T.H.G.; Hoang, X.B. User Privacy Risk Analysis within Website Privacy Policies. In Proceedings of the International Conference on Multimedia Analysis and Pattern Recognition (MAPR 2024), Da Nang, Vietnam, 15–16 August 2024; pp. 1–6. [CrossRef]
128. Jain, A.; Rodriguez, D.; del Alamo, J.M.; Sadeh, N. ATLAS: Automatically Detecting Discrepancies Between Privacy Policies and Privacy Labels. In Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroSPW 2023), Delft, The Netherlands, 3–7 July 2023; pp. 94–107. [CrossRef]
129. Kulkarni, M.; Naik, H.; Bharathi, S.V. Textual Analysis of Privacy Policies to Understand the Effect of GDPR. In Proceedings of the 2nd International Conference on Futuristic Technologies (INCOFT 2023), Belagavi, India, 24–26 November 2023; pp. 1–5. [CrossRef]
130. Luo, H.; Sun, Z.; Sun, Y.; Li, A.; Wang, B.; Cao, J.; Niu, B. SmartCircles: A Benefit-Evaluation-Based Privacy Policy Recommender for Customized Photo Sharing. In Proceedings of the IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta 2022), Haikou, China, 15–18 December 2022; pp. 2340–2347. [CrossRef]
131. Wang, T.; Hayes, C.; Chen, C.; Bashir, M. An Analysis of Mobile Gaming Apps' Privacy Policies. In Proceedings of the IEEE Games, Entertainment, Media Conference (GEM 2022), St. Michael, Barbados, 27–30 November 2022; pp. 1–6. [CrossRef]
132. Ibdah, D.; Lachtar, N.; Raparthi, S.M.; Bacha, A. "Why Should I Read the Privacy Policy, I Just Need the Service": A Study on Attitudes and Perceptions Toward Privacy Policies. *IEEE Access* **2021**, *9*, 166465–166487. [CrossRef]
133. Anikeev, M.; Shulman, H.; Simo, H. Privacy Policies of Mobile Apps—A Usability Study. In Proceedings of the IEEE INFOCOM 2021—IEEE Conference on Computer Communications Workshops, Vancouver, BC, Canada, 10–13 May 2021; pp. 1–2. [CrossRef]
134. Kuznetzov, M.; Novikova, E. Towards application of text mining techniques to the analysis of the privacy policies. In Proceedings of the 10th Mediterranean Conference on Embedded Computing (MECO 2021), Budva, Montenegro, 7–10 June 2021; pp. 1–4. [CrossRef]

135. Javed, Y.; Salehin, K.M.; Shehab, M. A Study of South Asian Websites on Privacy Compliance. *IEEE Access* **2020**, *8*, 156067–156083. [CrossRef]
136. Harkous, H.; Fawaz, K.; Lebret, R.; Schaub, F.; Shin, K.G.; Aberer, K. Polisis: Automated analysis and presentation of privacy policies using deep learning. In Proceedings of the 27th Conference on Security Symposium (USENIX 2018), SEC'18, Baltimore, MD, USA, 15–17 August 2018; pp. 531–548.
137. Kaur, J.; Dara, R.A.; Obimbo, C.; Song, F.; Menard, K. A comprehensive keyword analysis of online privacy policies. *Inf. Secur. J.* **2018**, *27*, 260–275. [CrossRef]
138. Wilson, S.; Schaub, F.; Liu, F.; Sathyendra, K.M.; Smullen, D.; Zimmeck, S.; Ramanath, R.; Story, P.; Liu, F.; Sadeh, N.; et al. Analyzing Privacy Policies at Scale: From Crowdsourcing to Automated Annotations. *ACM Trans. Web* **2018**, *13*, 1–29. [CrossRef]
139. Gonçalves de Pontes, D.R.; Zorzo, S.D. PPMark: An architecture to generate privacy labels using TF-IDF techniques and the rabin karp algorithm. *Adv. Intell. Syst. Comput.* **2016**, *448*, 1029–1040. [CrossRef]
140. Alohaly, M.; Takabi, H. If You Can't Measure It, You Can't Manage It: Towards Quantification of Privacy Policies. In Proceedings of the IEEE 2nd International Conference on Collaboration and Internet Computing (CIC 2016), Pittsburgh, PA, USA, 1–3 November 2016; pp. 539–545. [CrossRef]
141. Breaux, T.D.; Schaub, F. Scaling requirements extraction to the crowd: Experiments with privacy policies. In Proceedings of the IEEE 22nd International Requirements Engineering Conference (RE 2014), Karlskrona, Sweden, 25–29 August 2014; pp. 163–172. [CrossRef]
142. Savla, P.; Martino, L.D. Content Analysis of Privacy Policies for Health Social Networks. In Proceedings of the IEEE International Symposium on Policies for Distributed Systems and Networks (POLICY 2012), Chapel Hill, NC, USA, 16–18 July 2012; pp. 94–101. [CrossRef]
143. Benats, G.; Bandara, A.; Yu, Y.; Colin, J.N.; Nuseibeh, B. PrimAndroid: Privacy Policy Modelling and Analysis for Android Applications. In Proceedings of the IEEE International Symposium on Policies for Distributed Systems and Networks (POLICY 2011), Pisa, Italy, 6–8 June 2011; pp. 129–132. [CrossRef]
144. Earp, J.; Anton, A.; Aiman-Smith, L.; Stufflebeam, W. Examining Internet privacy policies within the context of user privacy values. *IEEE Trans. Eng. Manag.* **2005**, *52*, 227–237. [CrossRef]
145. Ding, J.; Li, X.; Guo, Y.; Yin, L.; Zhang, H. Process Calculus for Modeling and Quantifying Location Privacy. *Procedia Comput. Sci.* **2019**, *147*, 407–415. [CrossRef]
146. Viejo, A.; Sánchez, D. Enforcing transparent access to private content in social networks by means of automatic sanitization. *Expert Syst. Appl.* **2016**, *62*, 148–160. [CrossRef]
147. Cheng, H.-J.; Kumar, A. Process mining on noisy logs-Can log sanitization help to improve performance? *Decis. Support Syst.* **2015**, *79*, 138–149. [CrossRef]
148. Chen, X.; Pang, J. Protecting query privacy in location-based services. *GeoInformatica* **2014**, *18*, 95–133. [CrossRef]
149. Agarwal, N.; Kale, S.; Singh, K.; Thakurta, A. Improved differentially private and lazy online convex optimization: lower regret without smoothness requirements. In Proceedings of the 41st International Conference on Machine Learning (ICML 2024), JMLR.org, 2024, ICML'24, Vienna, Austria, 21–27 July 2024.
150. Li, J.; Lu, L. A Novel Differentially Private Online Learning Algorithm for Group Lasso in Big Data. *IET Inf. Secur.* **2024**, 5553292. [CrossRef]
151. Chen, L.; Ding, X.; Zhou, P.; Jin, H. Distributed dynamic online learning with differential privacy via path-length measurement. *Inf. Sci.* **2023**, *630*, 135–157. [CrossRef]
152. Zhang, M.; Sun, Z.; Li, H.; Niu, B.; Li, F.; Zhang, Z.; Xie, Y.; Zheng, C. Go-Sharing: A Blockchain-Based Privacy-Preserving Framework for Cross-Social Network Photo Sharing. *IEEE Trans. Dependable Secur. Comput.* **2023**, *20*, 3572–3587. [CrossRef]
153. Song, Y.; Ding, L.; Liu, X.; Du, M. Differential Privacy Protection Algorithm Based on Zero Trust Architecture for Industrial Internet. In Proceedings of the IEEE 4th International Conference on Power, Intelligent Computing and Systems (ICPICS 2022), Shenyang, China, 29–31 July 2022; pp. 917–920. [CrossRef]
154. Li, D.; Yang, Q.; Yu, W.; An, D.; Zhang, Y.; Zhao, W. Towards Differential Privacy-Based Online Double Auction for Smart Grid. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 971–986. [CrossRef]
155. Han, M.; Li, L.; Xie, Y.; Wang, J.; Duan, Z.; Li, J.; Yan, M. Cognitive Approach for Location Privacy Protection. *IEEE Access* **2018**, *6*, 13466–13477. [CrossRef]
156. Chen, L.; Zhu, P. Preserving the privacy of social recommendation with a differentially private approach. In Proceedings of the IEEE International Conference on Smart City (SmartCity 2015), Chengdu, China, 19–21 December 2015; Volume 2015, pp. 780–785. [CrossRef]
157. Balsa, E.; Troncoso, C.; Diaz, C. A metric to evaluate interaction obfuscation in online social networks. *Int. J. Uncertain. Fuzziness Knowldege-Based Syst.* **2012**, *20*, 877–892. [CrossRef]
158. Jia, J. A Consumer Data Privacy Protection Model Based on Non-Parametric Statistics for Dynamic Data Publishing in e-Commerce Platforms. *Hightech Innov. J.* **2024**, *5*, 410–419. [CrossRef]

159. Simo, H.; Kreutzer, M. WAPITI—A Weighted Bayesian Method for Private Information Inference on Social Ego Networks. In Proceedings of the IEEE 23rd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom 2024), Sanya, China, 17–21 December 2024; pp. 2269–2280. [CrossRef]
160. Neumann, G.K.; Grace, pp.; Burns, D.; Surridge, M. Pseudonymization risk analysis in distributed systems. *J. Internet Serv. Appl.* **2019**, *10*, 1. [CrossRef]
161. Liu, Y.; Zhang, W.; Yu, N. Protecting Privacy in Shared Photos via Adversarial Examples Based Stealth. *Secur. Commun. Netw.* **2017**, *2017*, 1897438:1–1897438:15. [CrossRef]
162. Pensa, R.G.; Di Blasi, G. A privacy self-assessment framework for online social networks. *Expert Syst. Appl.* **2017**, *86*, 18–31. [CrossRef]
163. Cevallos-Salas, D.; Estrada-Jiménez, J.; Guamán, D.S.; Urquiza-Aguiar, L. Ransomware dynamics: Mitigating personal data exfiltration through the SCIRAS lens. *Comput. Secur.* **2025**, *157*, 104583. [CrossRef]
164. Peng, J.; Xiang, X.; Qin, J.; Tan, Y. Dual-branch networks for privacy-preserving cross-modal retrieval in cloud computing. *J. Supercomput.* **2024**, *81*, 127. [CrossRef]
165. Arévalo, I.; Salmeron, J.L. A Chaotic Maps-Based Privacy-Preserving Distributed Deep Learning for Incomplete and Non-IID Datasets. *IEEE Trans. Emerg. Top. Comput.* **2024**, *12*, 357–367. [CrossRef]
166. Kumar, S.; Wajeed, M.A.; Kunabeva, R.; Dwivedi, N.; Singhal, P.; Jamal, S.S.; Akwafo, R. Novel Method for Safeguarding Personal Health Record in Cloud Connection Using Deep Learning Models. *Comput. Intell. Neurosci.* **2022**, *2022*, 3564436. [CrossRef]
167. Ng, P.C.; Spachos, P.; Gregori, S.; Plataniotis, K.N. Epidemic Exposure Tracking With Wearables: A Machine Learning Approach to Contact Tracing. *IEEE Access* **2022**, *10*, 14134–14148. [CrossRef]
168. Hairab, B.I.; Elsayed, M.S.; Jurcut, A.D.; Azer, M.A.; Anomaly Detection Based on CNN and Regularization Techniques Against Zero-Day Attacks in IoT Networks. *IEEE Access.* **2022**, 12, 98427–98440. [CrossRef]
169. Hui, S.; Wang, Z.; Hou, X.; Wang, X.; Wang, H.; Li, Y.; Jin, D. Systematically Quantifying IoT Privacy Leakage in Mobile Networks. *IEEE Internet Things J.* **2021**, *8*, 7115–7125. [CrossRef]
170. Thiruppathy Kesavan, V.; Loheswaran, K. A hybrid combination of ls-svm and kpca with bat algorithm for intrusion detection. *Recent Adv. Comput. Sci. Commun.* **2021**, *14*, 512–520. [CrossRef]
171. Kumar, A.; Braud, T.; Kwon, Y.D.; Hui, P. Aquilis: Using contextual integrity for privacy protection on mobile devices. *ACM Interact. Mob. Wearable Ubiquitous Technol.* **2020**, *4*, 1–28. [CrossRef]
172. Li, T.C.; Hang, H.; Faloutsos, M.; Efstathopoulos, P. Trackadvisor: Taking back browsing privacy from third-party trackers. In *Passive and Active Measurement. PAM 2015*; Lecture Notes in Computer Science Springer: Cham, Switzerland, 2015; Volume 8995, pp. 277–289. [CrossRef]
173. Archana, G.; Goyal, R.; Kumar, K. Blockchain-Driven Optimized Chaotic Encryption Scheme for Medical Image Transmission in IoT-Edge Environment. *Int. J. Comput. Intell. Syst.* **2025**, *18*, 11. [CrossRef]
174. Dogra, A.; Alkhayyat, A.; Saxena, A.; Dixit, K.; Singh, I.; Singh, S. Enhancing Data Privacy and Security in Healthcare IoT Applications Through Edge Data Filtering and Encryption for Secure and Compliant Cloud Transmission. In *Intelligent Systems Design and Applications (ISDA 2024)*; Springer: Cham, Switzerland, 2024; pp. 362–368. [CrossRef]
175. Routis, G.; Dagas, P.; Roussaki, I. Enhancing Privacy in the Internet of Vehicles via Hyperelliptic Curve Cryptography. *Electronics* **2024**, *13*, 730. [CrossRef]
176. Shetty, N.P.; Muniyal, B.; Priyanshu, A.; Kumar, D.; Melroy Maben, L.; Agrawal, Y.; Natarajan, R.; Gunda, S.; Gupta, N. Protecting Your Online Persona: A Preferential Selective Encryption Approach for Enhanced Privacy in Tweets, Images, Memes, and Metadata. *IEEE Access* **2024**, *12*, 86403–86424. [CrossRef]
177. Perazzo, P.; La Manna, M.; Iemma, F. Post-Quantum Attribute-Based Encryption: Performance Evaluation and Improvement for Embedded Systems. In Proceedings of the Post-Quantum Attribute-Based Encryption: Performance Evaluation and Improvement for Embedded Systems (EWSN 2022), Linz, Austria, 3–5 October 2022; Volume 1, pp. 268–273.
178. Torkamandi, P.; Pajevic Kärkkäinen, L.; Ott, J. Characterizing Wi-Fi Probing Behavior for Privacy-Preserving Crowdsensing. In Proceedings of the International Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2022), Montreal, QC, Canada, 24–28 October 2022; Volume 1, pp. 203–212. [CrossRef]
179. Aghasian, E.; Garg, S.; Montgomery, J. An automated model to score the privacy of unstructured information—Social media case. *Comput. Secur.* **2020**, *92*, 101778. [CrossRef]
180. Babaghayou, M.; Labraoui, N.; Ari, A.A.A.; Ferrag, M.A.; Maglaras, L. The Impact of the Adversary's Eavesdropping Stations on the Location Privacy Level in Internet of Vehicles. In Proceedings of the 5th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM 2020), Corfu, Greece, 25–27 September 2020; Volume 1, pp. 1–6. [CrossRef]
181. Papadopoulos, P.; Snyder, P.; Athanasakis, D.; Livshits, B. Keeping out the Masses: Understanding the Popularity and Implications of Internet Paywalls. In Proceedings of the The World Web Conference (WWW 2020), Taipei, Taiwan, 20–24 April 2020; Volume 1, pp. 1433–1444. [CrossRef]

182. Hu, D.; Hu, X.; Jiang, W.; Zheng, S.; Zhao, Z.Q. Intelligent digital image firewall system for filtering privacy or sensitive images. *Cogn. Syst. Res.* **2019**, *53*, 85–97. [CrossRef]
183. Sjösten, A.; Van Acker, S.; Sabelfeld, A. Discovering Browser Extensions via Web Accessible Resources. In Proceedings of the 7th ACM Conference on Data and Application Security and Privacy (CODASPY 2017), Scottsdale, AZ, USA, 22–24 March 2017; Volume 1, pp. 329–336. [CrossRef]
184. Meng, W.; Lee, B.; Xing, X.; Lee, W. Trackmeornot: Enabling flexible control on web tracking. In Proceedings of the 25th International World Wide Web Conference (WWW 2016), Montréal, QC, Canada, 11–15 April 2016; Volume 1, pp. 99–109. [CrossRef]
185. FaizKhademi, A.; Zulkernine, M.; Weldemariam, K. FPGuard: Detection and prevention of browser fingerprinting. In *Data and Applications Security and Privacy XXIX. DBSec 2015*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2015; Volume 9149, pp. 293–308. [CrossRef]
186. Dong, X.; Tran, M.; Liang, Z.; Jiang, X. AdSentry: Comprehensive and flexible confinement of JavaScript-based advertisements. In Proceedings of the ACM International Conference Proceeding Series (ICPS 2011), Orlando, FL, USA, 5–9 December 2011; Volume 1, pp. 297–306. [CrossRef]
187. Han, X.; Yang, Y.; Wu, J.; Xiong, H. HyObscure: Hybrid Obscuring for Privacy-Preserving Data Publishing. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 3893–3905. [CrossRef]
188. I, K.K.; S, H.A.A. RE-DACT: Adaptive Redaction and Anonymization Tool Using Machine-Learning. In Proceedings of the 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS 2024), Gobichettipalayam, India, 12–13 December 2024; pp. 708–713. [CrossRef]
189. Djomo, R.; Djotio Ndie, T. Towards Preventing Neighborhood Attacks: Proposal of a New Anonymization's Approach for Social Networks Data. In *Big Data Technologies and Applications. BDTA WiCON 2020 2020*; Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering; Springer: Cham, Switzerland, 2021; Volume 371, pp. 195–208. [CrossRef]
190. Gómez-Boix, A.; Frey, D.; Bromberg, Y.D.; Baudry, B. A collaborative strategy for mitigating tracking through browser fingerprinting. In Proceedings of the ACM Conference on Computer and Communications Security (ACM CCS 2019), London, UK, 11 November 2019; Volume 1, pp. 67–78. [CrossRef]
191. Mao, J.; Tian, W.; Yang, Y.; Liu, J. An Efficient Social Attribute Inference Scheme Based on Social Links and Attribute Relevance. *IEEE Access* **2019**, *7*, 153074–153085. [CrossRef]
192. Xie, Y.; Zheng, M. A differentiated anonymity algorithm for social network privacy preservation. *Algorithms* **2016**, *9*, 85. [CrossRef]
193. Mano, M.; Ishikawa, Y. Anonymizing user location and profile information for privacy-aware mobile services. In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks (SIGSPATIAL 2010), San Jose, CA, USA, 2 November 2010; Volume 1, pp. 68–75. [CrossRef]
194. Wang, Y.; Chen, X.; Jin, Q.; Ma, J. LIP3: A lightweighted fine-grained privacy-preserving profile matching mechanism for mobile social networks in proximity. In *Algorithms and Architectures for Parallel Processing. ICA3PP 2015*; Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering; Springer: Cham, Switzerland, 2015; Volume 9532, pp. 166–176. [CrossRef]
195. Chen, Z.; Wu, L. Residential appliance DR energy management with electric privacy protection by online stochastic optimization. *IEEE Trans. Smart Grid* **2013**, *4*, 1861–1869. [CrossRef]
196. Cevallos, D.; Cevallos, F.; Bernal, I.; Mejía, D. Internet Bot para la Obtención y Análisis de Información Empleando el Motor de Búsqueda de Google. *Rev. Politec.* **2015**, *35*, 38–48.
197. Cevallos, D.; Cevallos, F.; Bernal, I.; Mejía, D. Menu Creator: Plug-in de Generación Automática de Menús para el IDE de Desarrollo de Aplicaciones Interactivas NCL Composer. In Proceedings of the Twelfth LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI 2014) "Excellence in Engineering To Enhance a Country's Productivity", Guayaquil, Ecuador, 22–24 July 2014; pp. 1–10. https://laccei.org/LACCEI2014-Guayaquil/RefereedPapers/RP159.pdf.
198. Cevallos Salas, D.F. Enhanced RC5 Algorithm Using Parallel Computing for Communication Networks. *Ing. Y Cienc.* **2019**, *15*, 103–125. [CrossRef]
199. Congress.gov. S.5218—Health Infrastructure Security and Accountability Act of 2024. Available online: https://www.congress.gov/bill/118th-congress/senate-bill/5218/text (accessed on 22 July 2025).
200. Congress.gov. S.1851—Healthcare Cybersecurity Act of 2025. Available online: https://www.congress.gov/bill/119th-congress/senate-bill/1851 (accessed on 22 July 2025).
201. U.S. Department of Health and Human Services. The HIPAA Privacy Rule. Available online: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html (accessed on 22 July 2025).
202. Cevallos-Salas, D.; Estrada-Jiménez, J.; Guamán, D.S.; Rodríguez, D.; Del Álamo, J.M. GPT vs human legal texts annotations: A comparative study with privacy policies. *Research Square.* **2024.** [CrossRef]

203. Cevallos, F. Digital news classification and punctuaction using machine learning and text mining techniques. *Appl. Comput. Sci.* **2024**, *20*, 24–42. [CrossRef]
204. Wu, X.; Zou, B.; Lu, C.; Wang, L.; Zhang, Y.; Wang, H. Dynamic Security Computing Framework With Zero Trust Based on Privacy Domain Prevention and Control Theory. *IEEE J. Sel. Areas Commun.* **2025**, *43*, 2266–2278. [CrossRef]
205. Hussein, D.H.; Ibnkahla, M. Privacy-Preserving Intelligent Intent-Based Network Slicing for IoT Systems. *IEEE Internet Things J.* **2025**, *12*, 27783–27807. [CrossRef]
206. Mussgnug, A.M. Technology as uncharted territory: Contextual integrity and the notion of AI as new ethical ground. *arXiv* **2025**, arXiv:2412.05130. [CrossRef]
207. Sezgin, A.; Boyacı, A. DecoyPot: A large language model-driven web API honeypot for realistic attacker engagement. *Comput. Secur.* **2025**, *154*, 104458. [CrossRef]