The end of privacy by obscurity? Revisiting implicit privacy assumptions in the design of information systems

Andreis G. M. Purim[1], **Heitor P. Nolla**[2]

[1]Institute of Computing. University of Campinas (UNICAMP)

[2]School of Technology. University of Campinas (UNICAMP)

andreis.purim@students.ic.unicamp.br, heitor.nolla@gmail.com

Abstract. *Many information systems (IS) operate under a de facto model of "privacy by obscurity," in which users manage risk by limiting disclosure and adjusting settings. This model assumes that friction, legal authorization, and social norms constrain misuse. Contemporary data environments challenge this assumption. Large-scale aggregation, cross-context linkage, and reindentification (using machine learning models) reduce the protective value of limited visibility. This paper conceptualizes privacy by obscurity as an system design and identifies its core assumptions, examine the mechanisms through which it operates, and analyze the conditions under which it weakens. We then propose a taxonomy of privacy paradigms that compares obscurity with privacy by legality, architecture, market, and accountability approaches. The taxonomy clarifies how responsibility and enforcement are structured in contemporary information systems.*

1. Problem Context and Concept Introduction

A common response to privacy concerns in digital environments is: "if you do not want others to see it, do not share it." This view aligns with what has been described as the *privacy self-management* paradigm, in which individuals are expected to manage disclosure through notice-and-consent and selective participation [Solove 2013]. Due to the technical conditions of the last decades, this expectation was partially plausible: data was often fragmented across organizations, and aggregation across contexts required non-trivial effort and cost, in such a way that limiting disclosure could reduce exposure in practice. However, technical advances since then has shown that such assumptions do not hold under modern data processing capabilities, where re-identification and cross-context inference are feasible even from ostensibly limited or anonymised data [Ohm 2009, Narayanan and Shmatikov 2008].

Moreover, in many contexts, participation in digital systems is increasingly tied to employment, healthcare, education, finance, and communication, which reduces the practical feasibility of avoiding the sharing of personal data [Acquisti et al. 2015][1]. Due to the pervasive integration of digital systems into everyday environments (including smartphones, wearables, IoT devices, and sensor infrastructures), data collection and

[1]At the time of writing, one illustrative development is the growing push by governments to require age verification on social media platforms, such as Brazilian Law nº 15.211/2025. While such measures may reduce exposure risks for minors, they also introduce an additional surface for the collection, centralisation, and potential leakage of sensitive personal data. This example illustrates how participation in digital systems can, in practice, entail additional forms of data disclosure rather than reducing them.

behavioural inference can occur independently of explicit or intentional user disclosure. In the prevailing platform business model, firms are incentivised to construct detailed user profiles, even when direct disclosure is limited, through mechanisms such as third-party tracking, data brokerage, profiling, re-identification, and cross-system integration [Zuboff 2019].

Therefore, the response of "not sharing" effectively assigns primary responsibility for privacy protection to individuals. Users are expected to manage exposure by limiting disclosure (adjusting privacy settings) and exercising caution in digital participation. This allocation of responsibility has been widely critiqued as placing unrealistic cognitive and informational burdens on users [Solove 2013, Acquisti et al. 2015]. We refer to the IS design, governance structures, and patterns of user acceptance that embed this expectation as *privacy by obscurity*[2]. In this study, we present this pattern as a distinct paradigm in IS and examine its underlying assumptions and limitations.

This paper argues that *privacy by obscurity* operates as a de facto IS paradigm (Section 2.1), that it can be formalised through a set of underlying assumptions (Section 2.2), and that recent technological developments reduce its effectiveness (Section 2.3). We further situate privacy by obscurity within a broader taxonomy of legal, architectural, market, and accountability-based approaches to privacy, focusing on how each allocates responsibility and enforceability (Section 3).

2. Privacy by Obscurity

2.1. Privacy by Obscurity as a De Facto Paradigm

To define a *de facto paradigm* in Information Systems, we propose at least three conditions: (1) it is a pattern that recurs across systems and contexts, (2) it establishes a stable allocation of responsibility (technically, socially, or legally) between actors, and (3) it is implemented through consistent design and governance mechanisms. The question, then, is whether observed privacy practices satisfy these conditions.

Prior work in Information Systems and related fields has examined how privacy is governed through a combination of individual responsibility, system design, socionormative acceptance, and institutional arrangements. A central line of research describes the dominant model as "privacy self-management" or "privacy as control", in which individuals are expected to manage exposure through notice-and-consent and selective disclosure [Solove 2013, Bamberger and Mulligan 2015, Cate 2006]. This model places primary responsibility for privacy on individuals, while organizations retain control over data collection, aggregation, and reuse.

In practice, this allocation of responsibility is implemented through recurring design and governance mechanisms. Here, consent serves as a primary basis for legitimacy: users accept privacy policies and terms of service that authorize broad categories of data processing, often including sharing with third parties [Cate 2006]. Refusal is formally possible but frequently costly when access to services is tied to essential

[2]We do not adopt Solove's notion of "privacy self-management" [Solove 2013] as our primary term because it emphasises user *agency* and *capability* in managing disclosure. In contrast, we believe that *privacy by obscurity* refers to a system-level pattern in which privacy is maintained, in practice, through limited visibility, fragmentation, and contextual separation of data, regardless of whether users actively manage their exposure.

forms of participation, limiting the effectiveness of consent as a granular control mechanism [Acquisti et al. 2015].

Moreover, common design patterns in IS reinforce this allocation of responsibility by shaping how users make decisions. For example, default settings and interface design strategies systematically influence user choices and increase acceptance rates [Acquisti et al. 2015, Mathur et al. 2019]. Cookie banners and similar interfaces often prioritise rapid agreement over informed choice [Mathur et al. 2019]. As a result, these design choices actively steer users toward acceptance while preserving formal user agency. This steering effect is reflected in observed user behaviour. Empirical studies document forms of privacy fatigue and resignation, where repeated exposure to complex consent interactions weakens the relationship between stated preferences and observed decisions [Acquisti et al. 2015]. Under these conditions, acceptance becomes routine rather than deliberative. Individuals remain responsible for managing their exposure, while lacking visibility into how data is aggregated, inferred, and reused [Ohm 2009]. In other words, they assume the responsibility over their privacy but lose sovereingty or autonomy over their data.

Taken together, we defend that these mechanisms and behavioural outcomes satisfy the conditions outlined above. The same allocation of responsibility, the same design mechanisms, and the same behavioural responses recur across platforms and interfaces, *independently of specific implementations*. Privacy is therefore consistently treated as something individuals manage by controlling visibility, while organizations retain discretion over downstream processing. We therefore characterise *privacy by obscurity* as a *de facto paradigm* in Information Systems.

2.2. Core Assumptions and Mechanisms of Privacy by Obscurity

The previous section established privacy by obscurity as a recurring paradigm in Information Systems, now we identify the assumptions that support this paradigm and examine how they are realised in system design and governance. We distinguish three core assumptions[3]: (1) visibility as control, (2) consent as (persistent) legitimacy, and (3) feasible self-management. Together, these assumptions imply that privacy is achieved through reduced exposure and individual responsibility. Moreover, we believe this reasoning depends on two conditions: (i) sufficient friction in aggregating data across contexts, and (ii) limited capacity to infer sensitive attributes from partial signals. Under these conditions, limiting visibility can act as a practical form of control.

(1) Visibility/Obscurity as control. The first assumption is that reduced visibility (*obscurity*) yields protection. Information that is less accessible, less searchable, or less prominent is treated as functionally private. Earlier information environments supported this assumption: limited indexing and high aggregation costs constrained large-scale reuse. The notion of *practical obscurity* captured this condition, where information could be legally public but difficult to assemble at scale [Solove 2013]. In contemporary systems, visibility is primarily managed at the interface level. Functionally, platforms distinguish between "public", "friends-only", and "private" content, while backend infrastructures permit aggregation, internal reuse, and automated collection of accessible data. Authentication gates and rate limits regulate access volume.

[3]This set is not intended to be exhaustive, but to capture the core assumptions necessary to characterise the paradigm.

(2) Consent as legitimacy. The second assumption is that consent provides a sufficient basis for legitimate data processing. Agreement to terms of service and privacy policies is treated as authorisation for collection, retention, and reuse. These agreements typically cover broad categories of processing, including sharing with third parties. Refusal is formally possible, but may restrict access to services tied to participation in economic or social activities [Acquisti et al. 2015]. Within this framework, consent operates as the primary mechanism through which data practices are authorised. Legitimacy is established through acceptance of terms, and subsequent data processing is understood as grounded in this initial agreement [Solove 2013, Acquisti et al. 2015]. Moreover, consent is treated as temporally persistent: a single act of agreement is taken to authorise ongoing and future data processing.

(3) Self-management as feasibility. The third assumption is that individuals can manage privacy through selective disclosure and behavioural choices. This presumes that users can evaluate trade-offs and anticipate the consequences of sharing data. Privacy management is therefore framed as an individual task, carried out through decisions about what to share, when to share, and under which settings. In contemporary systems, this model of control operates across multiple data sources. Data collection includes explicit inputs, such as form entries and user-generated content, as well as implicit signals, such as telemetry, metadata, and behavioural traces captured across devices and services [Zuboff 2019]. These data sources contribute to a broader informational context in which individual disclosure decisions are interpreted.

These assumptions are reflected in recurring design and governance mechanisms: first, *interface defaults* frequently permit broader data collection or visibility, requiring users to actively restrict exposure [Acquisti et al. 2015, Mathur et al. 2019]. Second, *procedural consent structures* condition access to services on acceptance of standardised terms that authorise wide-ranging processing [Solove 2013, Bamberger and Mulligan 2015]. Third, *fragmented governance and data lineage opacity* limit visibility into how data moves across systems, complicating oversight [Ohm 2009]. Fourth, *emphasis on access control* prioritises authentication and access restriction, focusing on regulating access rather than downstream use. Together, these assumptions and mechanisms describe a consistent approach: privacy is managed by controlling visibility and individual behaviour, while data collection and processing remain broadly permitted within authorised systems.

2.3. Structural Limits of Privacy by Obscurity

The assumptions described above depend on conditions that no longer hold in contemporary systems. Increases in data availability, storage capacity, and computational power have reduced the cost of aggregation, while advances in statistical modelling and machine learning have expanded the capacity to infer sensitive attributes from partial signals. Under these conditions, limiting visibility does not provide reliable control over personal information.

First, the assumption of *visibility as control* breaks under large-scale aggregation and re-identification. Data that is individually non-sensitive can become identifying when combined with other sources. Early work showed that a large fraction of individuals can be uniquely identified using only a small number of quasi-identifiers such as ZIP code, birth date, and gender [Sweeney 2000]. Subsequent studies demonstrated that

anonymized datasets can be re-identified by linking them with auxiliary data, as in the case of the Netflix dataset [Narayanan and Shmatikov 2008]. More generally, empirical results show that human mobility traces, transaction records, and online behaviour are highly unique, even when partially observed [de Montjoye et al. 2013]. These findings indicate that visibility and identifiability are not equivalent: restricting access to data does not prevent individuals from being identified once data is combined.

Second, the assumption of *consent as (persistent) legitimacy* weakens when data processing extends beyond the context of collection. Data is routinely combined, transformed and/or transferred across organizations after the initial agreement. Consent provides a formal basis for authorisation, but it does not constrain how data is used once collected. In practice, a single act of acceptance enables multiple downstream uses, including aggregation, inference, and sharing with third parties. As a result, consent legitimises data flows without limiting their scope or effects.

Third, the assumption of *feasible self-management* fails under conditions of continuous data collection and inference. Individuals are expected to manage privacy through disclosure decisions, yet many data flows occur independently of explicit user input. At the same time, inference techniques extract sensitive attributes from behavioural data without requiring direct disclosure. For example, studies have shown that personality traits, political preferences, and other personal characteristics can be inferred from digital traces such as social media activity [Kosinski et al. 2013]. Effective control would therefore require users to anticipate how data from multiple sources may be combined and analysed, which is not practically achievable.

These limitations are also observable in deployed system: biometric identification techniques can recognize individuals from indirect signals such as gait or behaviour, including in the background of images or video - and such identification does not require explicit participation. At the same time, widespread recording in public and semi-public spaces enables the large-scale collection of these signals. These conditions extend data collection beyond deliberate disclosure and reduce the effectiveness of visibility-based control. Taken together, these developments show that privacy by obscurity fails at the level of its core assumptions. When aggregation is inexpensive, inference is reliable, and data collection is continuous, limiting visibility does not prevent identification, profiling, or secondary use. Privacy can no longer be maintained by reducing exposure alone.

3. A Taxonomy of Privacy Paradigms

To situate privacy by obscurity within a broader design space, we introduce a tentative taxonomy of privacy paradigms. Each paradigm describes a distinct way in which privacy constraints are produced, enforced, and allocated across actors. In practice, these paradigms do not operate in isolation, as contemporary systems combine multiple approaches, often within the same data pipeline - thus, our hypothesis is that they constitute different perspectives.

We differentiate these paradigms along four dimensions:

- **Enforcement mechanism:** How privacy constraints are operationalised (e.g., friction, legal authorisation, architectural constraint, market signalling, audit).
- **Primary locus of responsibility:** The actor expected to uphold privacy (user, firm, system designer, regulator).

- **Implicit trust condition:** The background assumption required for the paradigm to function (e.g., rational users, compliant firms, effective enforcement, competitive discipline).
- **Typical failure mode:** The condition under which the paradigm loses effectiveness.

Table 1. Privacy Paradigms in Contemporary Information Systems

Paradigm	Enforcement Mechanism	Responsible Actor	Trust Assumption	Typical Failure Mode
Privacy by Obscurity	Limited discoverability, contextual separation, friction to aggregation	Individual user	Data remains difficult to aggregate or interpret	Linkage, inference, persistent identifiers, cross-context integration
Privacy by Legality	Formal authorisation, contractual limitation, regulatory compliance	Firm + legal system	Terms are meaningful; compliance is monitored	Cognitive overload, vague drafting, weak enforcement, symbolic compliance
Privacy by Architecture	Ex ante technical constraints (e.g., access controls, minimisation, encryption)	System designer	Constraints are correctly implemented and not bypassed	Implementation gaps, purpose drift, internal override, adversarial circumvention
Privacy by Market	Competitive differentiation, reputational signalling, user switching	Private firm	Users can observe differences and switch providers	Information asymmetry, lock-in, network effects, externalities
Privacy by Accountability	Logging, auditability, traceability, sanction mechanisms	Organisation + regulator	Detection and penalties deter misuse	Audit gaps, fragmented oversight, limited transparency

This taxonomy[4] indicates that privacy by obscurity is not replaced by alternative paradigms, but often persists within them (as an asusmption). While legality, architecture, market mechanisms, and accountability introduce additional forms of constraint, they do not eliminate reliance on exposure-based control. *Privacy by legality* authorises data processing through consent and regulation, but often leaves downstream aggregation and inference unconstrained. *Privacy by architecture* restricts access through technical means, yet typically focuses on entry points rather than controlling how data is used after collection. *Privacy by market* assumes that users can observe and respond to differences in privacy practices, which depends on visibility that is often limited in practice. Privacy by accountability introduces ex post controls through auditing and sanction, but does not prevent data collection or inference from occurring.

As a result, these paradigms often coexist with (rather than replace) privacy by obscurity. Exposure-based control remains a background assumption, particularly where

[4]The labels used for these paradigms are tentative and intended as analytical conveniences rather than fixed or standardized terminology.

stronger constraints are partial, unevenly enforced, or limited in scope. Thus privacy by obscurity not as an outdated model and remains as a persistent component of contemporary systems. Its role shifts from a primary mechanism of protection to a residual logic that fills gaps left by other approaches.

4. Conclusion and Remaining Gaps

This paper characterizes privacy by obscurity as a de facto paradigm in information systems. We identify three core assumptions (visibility as control, consent as legitimacy, and self-management as feasibility) and shows how they are implemented through recurring design and governance mechanisms. These assumptions rely on conditions of limited aggregation and inference. In contemporary systems, these conditions no longer hold.

Our analysis presents privacy by obscurity alongside legal, architectural, market, and accountability-based approaches, highlighting differences in how responsibility and enforcement are structured. The central implication is that privacy cannot be evaluated in terms of exposure alone, but must be understood in terms of how data flows are constrained and how those constraints are enforced.

However, this argument is conceptual and leaves several points open to empirical validation. First, the prevalence of exposure-based reasoning in user decision-making has not been directly measured. Second, the shift in technical conditions (particularly the scale of data collection and inference) has not been quantified in a unified way. Third, the effectiveness of exposure-based controls in reducing identification or inference risk remains unclear in practice. Fourth, the role of social acceptance in sustaining exposure-based models is not well understood.

Addressing these points would strengthen the empirical grounding of the analysis, but does not alter its central claim: under current technical and economic conditions, *limiting visibility is not a reliable mechanism for protecting personal information.*

Acknowledgments

The authors acknowledge the use of Large Language Model (LLM) tools (ChatGPT and Grammarly) for language revision and editorial assistance. The research, development, analysis, and conclusions presented in this work are entirely those of the authors.

References

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221):509–514.

Bamberger, K. A. and Mulligan, D. K. (2015). *Privacy on the ground: driving corporate behavior in the United States and Europe*. MIT Press.

Cate, F. (2006). The failure of fair information practice principles. *Consumer Protection in the Age of the Information Economy*.

de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., and Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3(1).

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805.

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., and Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–32.

Narayanan, A. and Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, page 111–125. IEEE.

Ohm, P. (2009). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA l. Rev.*, 57:1701.

Solove, D. J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard law review*, 126(7):1880–1903.

Sweeney, L. (2000). Simple demographics often identify people uniquely.

Zuboff, S. (2019). *The age of surveillance capitalism*. PublicAffairs, New York, NY.

Biography

Andreis G. M. Purim (Andrejs G. M.-Puriņš) is a computer engineer and researcher in cybersecurity, cryptography, and artificial intelligence. He holds dual bachelor's degrees in Computer Engineering and Computer Science from the University of Campinas, as well a Diplôme d'Ingénieur (master's level) from the École Centrale de Lille and a master's degree in Law, Economics, and Management (Droit, Economie et Gestion) from the Université de Lille.

Heitor P. Nolla is an undergraduate student in Information Systems at the University of Campinas (UNICAMP), with a focus on machine learning, computer vision, and biometrics. His research involves generative models and deep neural networks applied to the synthesis and analysis of biometric data