

The Right Not to Be Inferred: Profiling, Disinformation, and the Limits of the Right to Data Protection

Josué Sá Fama^{ID}

Data Protection Officer, Instituto Atlântico, Fortaleza, Brazil

Email: josue_fama@atlantico.com.br

How to cite this paper: Fama, J. S. (2026). The Right Not to Be Inferred: Profiling, Disinformation, and the Limits of the Right to Data Protection. *Beijing Law Review*, 17, 729-745.

<https://doi.org/10.4236/blr.2026.173038>

Received: April 27, 2026

Accepted: June 29, 2026

Published: July 2, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Contemporary data protection regimes were designed to regulate the collection and processing of personal data, but increasingly fail to address epistemic and behavioral harms produced by algorithmic inferences. This article argues that profiling and micro-targeting of misinformation exploit a structural gap in frameworks, which prioritize control over inputs (collected data) while leaving inferential outputs (inferred attributes, profiles, models, and predictions) relatively ungoverned. By articulating the political economy of surveillance capitalism, cognitive vulnerabilities, and a legal analysis of the GDPR, LGPD, and Digital Services Act (DSA), it demonstrates how algorithmic inference erodes informational self-determination and, consequently, democratic autonomy. Empirically, 1) the analysis of 856 GDPR fines indicates a predominance of organizational and technical violations centered on early stages of the data lifecycle, with less regulatory traction on profiling and inference harms (Sae-mann et al., 2022); 2) large-scale evidence shows that fake news spreads faster and reaches more people than true news (Vosoughi, Roy, & Aral, 2018); and 3) political advertising data map the global expansion of microtargeting (Votta, Kruschinski, & Hove, 2024). Finally, it is proposed that a normative limit to inference be recognized, formulated as the right not to be inferred (under specific conditions and categories) as a necessary evolution of data protection law in societies mediated by artificial intelligence.

Keywords

Data Protection, Algorithmic Inference Profiling, Informational Self-Determination, Democracy

1. Introduction

Digital society is undergoing the ontological transformation described by Floridi

as the “Fourth Revolution,” in which human agency begins to operate within the infosphere, dissolving the boundary between online and offline in an “onlife” condition (Floridi, 2014). In this environment, public decisions and social coordination depend on infrastructures that process, rank, and recommend information on a large scale.

The promise of an open and deliberative public sphere has been strained by an epistemic crisis marked by polarization, disinformation, and manipulation. To address the phenomenon rigorously, the article avoids the politically captured term “fake news” and adopts the distinction between misinformation, disinformation, and malinformation, consolidated by Wardle and Derakhshan (2017), as a taxonomy of “informational disorder” (Wardle & Derakhshan, 2017).

1.1. Research Problem

The central legal problem is not simply about “data being collected.” The core issue shifts to the power to infer, which is the ability to produce attributes and profiles not provided by the data subject and to use these outputs to target information and shape behavior, including in the exercise of citizenship.

1.2. Research Question

To what extent (and through what mechanisms) do the LGPD (Brazilian General Data Protection Law) and related regimes (such as the GDPR), by primarily regulating the collection and processing of personal data, leave the production and use of algorithmic inferences (outputs) unchecked, allowing for profiling and microtargeting of misinformation with epistemic, behavioral, and democratic impacts, and what normative parameters would allow for the recognition of a “right not to be inferred” as a regulatory response?

1.3. Central Hypothesis

H1: Data protection regimes centered on inputs (collection/use) are structurally insufficient to protect informational self-determination and democratic autonomy because they do not adequately govern the production and use of inferences (outputs), allowing profiling and microtargeting to operate in a largely permissive manner even when they produce epistemic harm, defined here as the degradation of an individual’s capacity to form accurate beliefs and exercise rational judgment, including through exposure to algorithmically curated disinformation, and behavioral harm, defined as the manipulation of conduct, preferences, or civic participation through inference-driven targeting that bypasses conscious deliberation, with cumulative consequences for democratic autonomy, understood as the collective capacity of citizens to form political will under conditions of epistemic fairness, transparency of persuasion, and equal access to an unmanipulated informational environment.

1.4. Complementary Hypotheses

H2 (enforcement): Sanctioning and guidance practices tend to focus on organiza-

tional/technical failures in the data cycle (legal basis, security, minimization), with less focus on inferential harm and profiling, reinforcing the input-output gap (Saemann et al., 2022).

H3 (democracy): Even when the average persuasive effectiveness of microtargeting is “mixed,” its capacity for opaque differential treatment of population segments (mobilization/demobilization and contradictory messages) is already sufficient to generate significant democratic risk, as it reduces public accountability and fragments the informational environment, thereby undermining democratic autonomy as defined above.

1.5. Method and Evidence

The article combines: 1) legal-normative analysis (GDPR, LGPD, and DSA); 2) theoretical review (surveillance capitalism, cognitive vulnerabilities, manipulation); and 3) empirical triangulation with three pieces of evidence: GDPR fines (Saemann et al., 2022), the spread of true and false news (Vosoughi, Roy, & Aral, 2018), and large-scale political microtargeting (Votta, Kruschinski, & Hove, 2024), complemented by sectoral and experimental studies (Zarouali, Dobber, & Schreuder, 2023).

1.6. Rationale for Empirical Source Selection

The three primary empirical sources were selected because each addresses a distinct causal layer of the article’s argument. First, the GDPR fine dataset (Saemann et al., 2022) operationalizes Hypothesis H2 by providing systematic, quantitative evidence of enforcement patterns: it establishes what regulators actually sanction, revealing the input-output asymmetry at the institutional level. Second, the study of true and false news diffusion (Vosoughi, Roy, & Aral, 2018) provides large-scale causal evidence that the platform architecture structurally advantages misinformation, establishing the systemic epistemic harm that inference-driven engagement amplifies. Third, the transnational microtargeting dataset (Votta, Kruschinski, & Hove, 2024) documents the scale and geographic scope of political inference deployment, grounding the democratic risk claim in observable advertiser behavior rather than hypothetical capability. Together, the three sources triangulate across the enforcement gap (institutional), the harm mechanism (diffusion), and the deployment reality (political targeting), making the normative proposal for a right not to be inferred empirically tractable rather than purely speculative.

2. Surveillance Capitalism, the Attention Economy, and the Architecture of Disinformation

The hypothesis of this work presupposes that informational disorder is less an “accident” and more a systemic product of an economic architecture: monetization of attention via data extraction and inference generation (Zuboff, 2019). Platforms are not neutral channels; they are designed to maximize engagement, and inference is the engine that makes engagement predictable, testable, and monetiz-

able.

2.1. Economic Scale of the Model (Inference as Infrastructure)

The scale of the behavior-driven advertising market suggests why inference is not peripheral: it is the foundation of the model. Meta reports annual revenue of approximately US\$164.5 billion in 2024 and advertising revenue of approximately US\$160.6 billion in the same year (Meta Platforms Inc, 2025). In the same period, Google advertising revenue reaches approximately US\$264.6 billion in 2024 (Alphabet Inc., 2026). Furthermore, market reports estimate the global data broker market to be in the hundreds of billions, with projections of significant growth (Grand View Research, 2026).

Table 1. Economic scale associated with advertising and data intermediation (selection).

Entity	Year	Metric	Value (USD)	Source
Meta	2024	Total revenue	164,501 billion	(Meta Platforms Inc, 2025)
Meta	2024	Advertising revenue	160,633 billion	(Meta Platforms Inc, 2025)
Alphabet/Google	2024	Google advertising	264,590 billion	(Alphabet Inc., 2026)

As shown in Table 1, in a normative reading, these numbers do not “prove” harm per se. They demonstrate structural incentives; that is, if income depends on personalized advertising and targeting, then the production of inferences (profiles, predispositions, vulnerabilities) tends to be constant, amplified, and optimized.

2.2. Misinformation as a “Feature,” Not a “Bug”

Engagement-centric design favors emotionally activating content (fear, anger, outrage), which tends to be more shareable. Classic large-scale evidence shows that fake news spreads faster and reaches more people than true news (Vosoughi, Roy, & Aral, 2018). Thus, misinformation ceases to be merely “wrong content” and becomes functional content within an architecture that rewards engagement. This dynamic constitutes a direct mechanism of epistemic harm: the systematic degradation of the informational environment reduces citizens’ capacity to form accurate beliefs, with downstream effects on democratic autonomy.

As shown in Table 2, large-scale studies on rumor cascades on the former Twitter (now platform X) in 2006 and 2007 show that falsehoods spread faster than truths. This means that systems optimized for engagement and content with novelty/arousal have a structural advantage, which favors misinformation as a systemic phenomenon.

3. Profiling, Microtargeting, and Cognitive Vulnerability

To understand how profiling operates within the data protection ecosystem, it is

helpful to visualize the complete flow from the moment the user consents to the final harm produced (Figure 1):

Table 2. Comparison of diffusion: Truthfulness vs. Falsehood.

Analysis Variable	Falsehood Performance	Comparison with Truth
Speed	Significantly faster diffusion	Truth takes six times longer to reach 1500 people.
Reach (Depth)	Reaches much deeper cascade levels.	Truth rarely goes beyond 10 levels of depth.
Amplitude (Width)	Wider cascades (greater virality)	Falsehood: 1000 - 100,000 people; truth is rarely > 1000
Psychological Factor	Surprise and disgust	Sadness, anticipation, and trust
Users	Spread by humans (not just bots)	Humans are more likely to retweet false information.

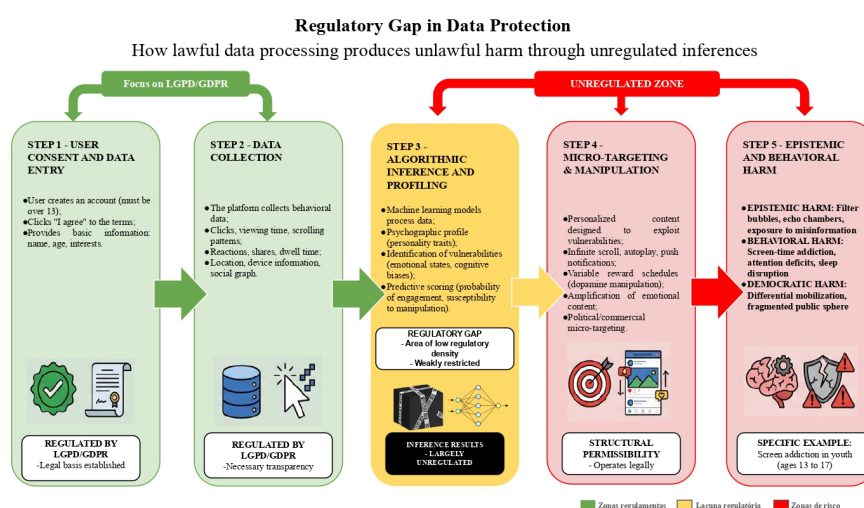


Figure 1. Regulatory gap in data protection—from consent to harm.

Figure Labels: The diagram illustrates the five-step process from user consent to epistemic and behavioral harms. The green zones (Steps 1 - 2) represent areas heavily regulated by GDPR/LGPD through consent, transparency, and data minimization requirements. The yellow zone (Step 3) identifies the critical regulatory gap: it is at this stage that profiling occurs, the algorithmic transformation of behavioral data into inferred psychological profiles about vulnerabilities, preferences, and susceptibilities. Although GDPR Article 22 and Article 4(4) mention profiling, the restrictions are weak, and the governance of inference remains in a legal grey area. The red zones (Steps 4 - 5) show how unregulated inference allows manipulative microtargeting practices (Step 4) that produce measurable behavioral and epistemic harms (Step 5), with aggregate consequences for democratic

autonomy. The diagram highlights the structural asymmetry: the law strongly regulates the inputs (data collection) but weakly governs the inferential transformation and leaves the outputs (harm) largely outside the scope of data protection.

The Critical Moment of Profiling: Step 3

Profiling: defined in the GDPR (Article 4(4)) as “any form of automated processing of personal data consisting of using that personal data to evaluate certain personal aspects relating to a natural person,” occurs precisely in Step 3 of the diagram. In this stage, platforms use machine learning models to transform observable behavioral data (time spent, click patterns, interaction history) into inferred attributes not provided by the data subject.

- Personality traits (openness, neuroticism, extraversion)
- Emotional vulnerabilities (susceptibility to anxiety, social comparison, fear)
- Sensitivity to rewards and intermittent reinforcement schedules
- Political and ideological predispositions
- Attention patterns and indicators of habit formation

This transformation is at the heart of the regulatory gap. While the collection of behavioral data (Step 2) is governed by principles of minimization, transparency, and purpose (GDPR Articles 5, 13-14), and initial consent (Step 1) is governed by validity requirements (GDPR Articles 6-8), the production of inferences operates with weak constraints. Article 22 of the GDPR provides limited rights regarding “automated individual decision-making,” but only when such processing “produces legal effects concerning [the data subject] or similarly significantly affects him or her.” Courts and regulators have interpreted this narrowly, leaving profiling that “merely” shapes exposure to information, emotional states, or behavioral tendencies outside robust protection.

The result: platforms can build sophisticated profiles of psychological vulnerability (Step 3), use them for manipulative microtargeting (Step 4), and produce measurable behavioral and epistemic harm (Step 5) through a process that is largely legitimate at each individual step, but structurally harmful in the aggregate. As empirically demonstrated in the Cambridge Analytica case, which collected 87 million Facebook profiles and used OCEAN psychographic modeling to infer personality traits from behavioral data (Hu, 2020), profiling can operate on a massive scale with minimal regulatory restraint. The identified violation was unauthorized access to the data (Step 2), not the inferential practices themselves (Step 3). This is the input-output asymmetry that structures the central legal problem of this article.

3.1. Profiling as the Production of “Knowledge” about the Data Subject

Profiling converts observable signals (clicks, dwell time, network of connections) into inferred attributes (preferences, likely political orientation, personality traits, susceptibilities). Here, a legal twist emerges: the information that most affects the

data subject may not be what he provided, but what was inferred about him (Hildebrandt, 2015; Pasquale, 2015).

3.2. Political Microtargeting at Scale

Recent evidence maps political microtargeting on a transnational scale. A study based on advertising libraries reports tens of thousands of political advertisers and millions of targeted ads across dozens of countries and elections (Votta, Kruschinski, & Hove, 2024). In a focus group study, political ads in Spanish elections show the systematic use of demographic, geographic, and interest criteria for distribution optimization.

Table 3. Political microtargeting.

Dataset	Metric	Value	Period	Source
Ad Library (Meta)	Advertisers	54,000	2020-2022	(Votta, Kruschinski, & Hove, 2024)
Ad Library (Meta)	Advertisements	2.5 million	2020-2022	(Votta, Kruschinski, & Hove, 2024)
Ad Library (Meta)	Countries/elections	95/113	2020-2022	(Votta, Kruschinski, & Hove, 2024)

As shown in Table 3, the mapping carried out on Meta platforms, covering 95 countries and 113 national elections between 2020 and 2022, shows that the practice is widespread and varies significantly according to the context and segmentation criteria used. The relevance of this phenomenon lies in the fact that microtargeting allows for differential persuasion, the sending of messages “in the dark,” and the consequent fragmentation of the public sphere, posing serious challenges to democratic autonomy by reducing the transparency of political conviction and fragmenting the shared informational environment necessary for deliberation.

3.3. Exploitation of Cognitive Vulnerabilities

Inference becomes dangerous when used to calibrate messages to predictable biases (confirmation, heuristics, loss aversion), reducing reflection and amplifying reactivity (Kahneman, 2011; Sunstein, 2017). This calibration is the primary mechanism of behavioral harm: by bypassing conscious deliberation and targeting inferred susceptibilities, microtargeting interferes with the volitional and epistemic preconditions for autonomous political action. Participatory misinformation, in which ordinary users also act as vectors of propagation, exacerbates the epistemic harm, as messages are “validated” by social signals (Wardle & Derakhshan, 2017).

3.4. Effectiveness of Microtargeting: Mixed Evidence and Persistent Democratic Risk

Empirical literature is not uniform; it indicates that the effectiveness of microtar-

getting is not an absolute rule. Some studies show modest or contextual effects and heterogeneous results (Zarouali, Dobber, & Schreuder, 2023). It is essential to distinguish between what these studies directly establish and what the article infers normatively. The cited experimental evidence directly demonstrates that: a) microtargeting can affect stated intention but does not consistently translate into observable behavior change; and b) congruence between inferred personality and ad susceptibility emerges through self-report but automated inference does not reliably replicate it (Zarouali, Dobber, & Schreuder, 2023). The article’s claim that microtargeting poses a democratic risk is a normative inference, drawn from the structural capacity for opaque differential treatment, rather than a finding directly established by the cited studies. This distinction is explicit: the democratic risk that matters here is independent of a stable average superiority, because microtargeting allows for opaque differential treatment (divergent messages for different audiences, potential suppression of participation by some groups, absence of public scrutiny of what was promised to whom), which is sufficient to threaten democratic autonomy even when individual persuasive effects are modest.

Table 4. Studies on the effectiveness of microtargeting.

Study	Design	Central finding	Source
Micro-targeted campaign	Field experiment	May affect intention but does not always translate into observable behavior.	(Zarouali et al., 2022)
Personality and susceptibility	Experiment	Congruence emerges through self-report; automated inference does not always replicate.	(Zarouali, Dobber, & Schreuder, 2023)

As shown in Table 4, the technique faces two major bottlenecks: the low accuracy of the algorithms for mapping our minds and the real difficulty of transforming digital attention into concrete human behavior.

4. The Structural Gap in Law: Input-Output Asymmetry (Data × Inferences)

Contemporary literature converges on the criticism that data protection law has been applied with an emphasis on controlling the collection and processing of personal data, but with weaker instruments to govern inferences, profiles, and models (Wachter & Mittelstadt, 2019). The result is an “input-output asymmetry”: what enters (data) is regulated with density, but what exits (inferences) is poorly regulated.

4.1. Evidence of Enforcement: What Results in a Fine (and What Doesn’t)

Analysis of 856 GDPR fines indicates a predominance of organizational (legal ba-

sis, governance, consent) and technical (security, infrastructure) problems, with a concentration in the initial phases of the data lifecycle (collection) (Saemann et al., 2022). In other words, the sanction tends to “see” traditional data lifecycle flaws more clearly than inferential damages.

Table 5. Penalty standards under the GDPR (total number: 856).

Category	Quantity (mentions)	Interpretation of the inferential gap	Source
Organizational problems	601	Strong emphasis on governance/inputs	(Saemann et al., 2022)
Technical problems	314	Impact on security/cycle architecture	(Saemann et al., 2022)

As shown in **Table 5**, the table makes it clear that compliance with the GDPR (and similar laws like the LGPD) fails much more due to bureaucracy, lack of processes, and human error than due to software obsolescence or sophisticated cyberattacks.

4.2. Why Do Inferences “Escape”? Three Recurring Frictions

1) Practical ambiguity exists regarding when inferences are treated as “personal data” and which rights are granted by inferred attributes (Wachter & Mittelstadt, 2019).

2) Technical opacity: even when there is formal transparency, challenging inferences may be unfeasible (Hildebrandt, 2015; Pasquale, 2015).

3) Trade secret/IP: models and logic can be protected, reducing effective access and auditability.

4) Recent peer-reviewed scholarship on inferential privacy: Two recent contributions reinforce and refine the input-output gap thesis. Wachter, Mittelstadt, and Russell (2021) demonstrate that the legal concept of “personal data” under the GDPR is structurally under-inclusive with respect to group-level inferences: platforms routinely generate behavioral predictions that are technically “non-personal” (derived from aggregate patterns) yet individually applied, creating a regime where the most consequential inferential outputs escape the categories that trigger data subject rights. Their analysis of “statistical discrimination” provides a doctrinal foundation for the enforcement gap documented by Saemann et al. (2022). Complementing this, Veale and Borgesius (2021) analyze the limits of Article 22 GDPR as an instrument for governing automated profiling, arguing that the “legal or similarly significant effects” threshold functionally exempts the vast majority of commercial profiling operations, precisely those producing epistemic harm through information shaping, from meaningful oversight. Both studies support Hypothesis H2 and are cited in Sections 5 and 6, where the doctrinal analysis of GDPR/LGPD provisions is developed.

5. Informational Self-Determination and Limits of the GDPR and LGPD

Informational self-determination presupposes that the individual can know, understand, and control the flow and use of information about themselves, and, crucially, challenge unfair representations (Rouvroy & Pouillet, 2009; Lynskey, 2015). In an inferential environment, the holder is continually reconstituted by predictions and classifications that they did not provide, generating ongoing epistemic harm in the sense that the data subject's capacity for self-representation and accurate belief formation is undermined by profiles they cannot access or contest.

5.1. GDPR: Art. 22 and the “Gap” between Profiling and Decision-Making

Article 22 of the GDPR focuses on automated decisions with legal or similar effects, which opens up grey areas: many profiling and microtargeting practices operate “below the threshold” or with “nominal human” involvement (WP29, 2018). EDPB guidelines on targeting show the effort to provide a framework, but do not eliminate the structural problem: regulating targeting without substantively governing the inference that feeds it limits regulatory effectiveness (EDPB, 2021). As Veale and Borgesius (2021) demonstrate, the Article 22 threshold is so narrowly interpreted that information shaping, the primary mechanism of behavioral harm in electoral contexts, falls entirely outside its protective scope, confirming the input-output asymmetry at the level of positive law.

5.2. LGPD: Convergence of the Challenge

The LGPD (Brazilian General Data Protection Law) provides for the review of automated decisions, but faces the same tension: the legal focus usually falls on legal bases, transparency, and procedural rights, while the governance of inferences remains more open, especially when the harm is diffuse or collective (Bioni, 2019). The recent Brazilian discussion on “reasonable inferences” suggests a path for interpretative evolution (de Miranda, 2022).

5.3. The “Fiction of Consent”

Even with consent, informational asymmetry (technical complexity, indispensability of services, decision fatigue, and dark patterns) compromises the ideal of self-determination via “privacy self-management” (Solove, 2013; Barocas & Nissenbaum, 2014). Consenting to data collection does not equate to consenting to future inferential outcomes that are still unknown.

5.4. Inferential Privacy and the Reasonable Inference Doctrine

Wachter, Mittelstadt, and Russell (2021) propose a right to reasonable inferences as a complementary doctrinal construct within the existing GDPR architecture, arguing that inferences should be subject to accuracy, necessity, and proportionality review even when they do not directly trigger Article 22. This is the closest

peer-reviewed analogue to the right not to be inferred proposed here and provides the principal doctrinal precedent for the proposal in Section 7. The distinction between the two formulations is deliberate: while Wachter et al. (2021) focus on the reasonableness of any given inference, the present article proposes a categorical prohibition regime for specific high-risk inference categories, going beyond reasonableness review to a justiciable right of non-subjection.

6. DSA and the Shift towards Systemic Risk: Progress, but Incomplete

The DSA introduces a systemic risk paradigm, including risks to democratic processes and civic discourse, and requires assessment and mitigation for very large platforms (European Union, 2022). This is an advance because it recognizes that informational harm is also collective. However, one point remains: systemic risk without substantive control of inference can turn into “auditing reports,” without shifting incentives for profiling and microtargeting.

DSA Systemic Risk Governance: Peer-Reviewed Assessment

Leerssen et al. (2023) provide the most detailed empirical assessment of the DSA’s systemic risk provisions, analyzing the gap between the obligation to conduct systemic risk assessments (Article 34) and the absence of substantive standards governing what those assessments must find or prohibit. Their central finding, that the DSA creates process obligations without inference prohibitions, empirically validates the concern raised in this section: accountability without substantive inferential governance is structurally insufficient to protect democratic autonomy. The authors argue that Article 34 systemic risk assessments, as currently structured, are unlikely to produce enforceable constraints on psychographic profiling or vulnerability-based political targeting absent complementary sector-specific rules. This finding strengthens the normative case for the Inference Impact Assessment (IIA) proposed in Section 7, which would operate as precisely such a complementary instrument.

7. Proposal—The Right Not to Be Inferred (Minimum Content and Operationalization)

The normative thesis of this article is that data protection needs to evolve from “control of data flow” to governance of the power of knowledge: whoever controls inference controls predictability and modulation capacity (Wachter & Mittelstadt, 2019).

The right not to be inferred is not a general prohibition of analytics. It is a normative limit directed at inferences that:

- Are of high impact (affecting rights, opportunities, reputation, informational exposure), or
- Are geared towards manipulation (especially in politics and civic information).

7.1. Definition of the Proposal

The Right Not to Be Inferred: the right of the data subject (and, in certain categories, of the community) not to have their attributes, profiles, or vulnerabilities inferred and used in high-impact contexts without a robust, auditable, and contestable justification—including cases of categorical prohibition.

Therefore, the focus is on:

- Inferences of cognitive susceptibility and vulnerability to persuasion,
- Psychographic profiling/susceptibility targeting in politics.
- Sensitive inferences that enable discrimination and steering,
- Inferences used to shape informational presentations on civic topics.

7.2. Legal Status of the Right Not to Be Inferred

The right not to be inferred is proposed here as a sector-specific autonomous right, operative primarily in high-risk political and civic contexts, rather than a reinterpretation of existing GDPR/LGPD provisions or a purely horizontal extension of current data subject rights. This classification matters doctrinally. As a reinterpretation, the right would be constrained by the text and case law of Articles 22 GDPR and its LGPD equivalent, which, as demonstrated above, are structurally limited. As a new autonomous right, it would require primary legislation and face the risk of regulatory dilution in transposition. As a sector-specific rule, it can be introduced through targeted regulation—analogue to the DSA’s Article 34 obligations or election-law frameworks—with bespoke enforcement mechanisms and a defined scope limited to political targeting, electoral advertising, and inference-driven civic information shaping. This is the approach advocated here, on the grounds that it is both legally achievable within existing constitutional frameworks for data protection (GDPR Recital 4; LGPD Art. 2) and sufficiently precise to avoid overbreadth objections. Accordingly, the right not to be inferred operates as a *lex specialis* within the data protection ecosystem, complementing rather than replacing the general regime.

7.3. Concrete Triggers for “High-Impact” Inferences

The following triggers are proposed to delimit the scope of the right not to be inferred, providing the legal certainty necessary for enforcement:

- 1) Inference of sensitive attributes (as defined in GDPR Art. 9 and LGPD Art. 11) from behavioral proxies, including political opinion, health, religion, sexual orientation, and ethnicity, without the data subject’s explicit and informed consent to the inferential act specifically (not merely to data collection).
- 2) Election-related or referendum-related targeting: any inference-driven advertisement, content recommendation, or information shaping directed at identifiable voter segments during an electoral period (defined as the 90 days preceding any national, regional, or municipal election).
- 3) Vulnerability-based civic persuasion: microtargeting that uses inferred psychological susceptibilities (neuroticism, anxiety propensity, low epistemic confi-

dence, grief, economic precarity) to calibrate civic or political messaging, regardless of the electoral period.

4) Behavioral inference producing legal or quasi-legal effects: any inference used in automated or semi-automated systems that produces consequences equivalent to those regulated by GDPR Art. 22 (creditworthiness, employment screening, insurance pricing, access to public services) even when formal “decision” elements are absent.

5) Large-scale psychographic modeling: the systematic construction of personality, emotion, or ideological profiles affecting more than a defined threshold of individuals (proposed: 10,000 data subjects within any 12-month period) for commercial or political purposes.

7.4. Four Normative-Operational Pillars

1) Justification *ex ante* (Inference Impact Assessment-IIA): the duty to justify the need, proportionality, validity/error of the inferential assessment, and mitigation. This is the controller’s burden.

2) Narrow (defensible) categorical prohibitions: inferences intended to exploit cognitive vulnerabilities for political manipulation; psychometrics for political microtargeting; sensitive inferences of low verifiability in asymmetrical contexts. Prohibition of inferences whose objective is to identify vulnerabilities for stealthy political manipulation.

3) Real contestability and access to inferential categories: access to inferred categories, effects, and means of contestation, with correction/elimination when applicable (not just “vague explanations”). That is, the right to know (at least by categories) what was inferred, for what purpose, and with what effects, with mechanisms to contest unreasonable inferences.

4) Democratic dimension (democratic collective safeguards): recognition that part of the damage is collective (fragmentation of public space, opacity of political conviction, erosion of the epistemic conditions of deliberation). That is, transparency and accountability that protect the exercise of citizenship and civic discourse beyond purely individual logic.

5) Implementation of the Inference Impact Assessment (IIA): The IIA is proposed as an *ex ante* procedural obligation structured as follows. Preparation: the IIA must be prepared by the controller (platform or political advertiser) prior to the deployment of any inference system falling within the high-impact triggers identified above. It must document: a) the inferential model’s purpose, training data, and error rates disaggregated by demographic group; b) the categories of attributes inferred; c) the targeting logic linking inferences to content delivery; and d) the proportionality justification, including whether less inference-intensive means were considered. Independent Data Protection Officers (DPOs) must co-sign the IIA and cannot be overridden by controllers on scope determinations. Review: IIAs must be filed with the competent supervisory authority (the national DPA or the ANPD in Brazil) at least 30 days before deployment. For election-

related targeting, review must be completed within 15 days, with interim suspension as the default pending clearance. Supervisory authorities may request third-party algorithmic audits as a condition of approval, funded by the controller. Corrective measures upon failure: if an IIA reveals that the proposed inference system cannot meet proportionality requirements, or if a post-deployment audit reveals undisclosed inferential practices, the following graduated remedies apply: (i) mandatory suspension of the inference function; (ii) deletion of inferred profiles and prohibition on re-inference for a defined period; (iii) administrative fines calculated as a percentage of global advertising revenue (analogous to GDPR Art. 83(5)), with enhanced penalties for election-period violations; and (iv) in cases of systematic or intentional violation, prohibition on political advertising for the platform for a defined electoral cycle. Remedies (i) and (ii) are immediate upon findings; (iii) and (iv) require adversarial proceedings with the controller.

Table 6. From the right to data protection to the governance of inference (summary).

Governance target	“Classic” data protection	Governance of inference (“right not to be inferred”)	Governance target
Central object	Legality of data processing	Inferential outputs + targeting	Central object
Key tools	Legal basis, purpose, consent/transparency, security	Impact assessments (IIA), auditability, non-profiling alternatives, high-risk point bans	Key tools
Damages addressed	Improper collection, secondary use, leaks	Epistemic harm (opaque profiles), behavioral steering/manipulation, democratic harm	Damages addressed
Accountability model	Individual rights + enforcement by DPA	Individual rights + collective safeguards (electoral integrity/systemic risk)	Accountability model

As shown in **Table 6**, the table indicates the transition from an era of Information Privacy (where the data is yours and you control who sees it) to an era of Cognitive/Behavioral Privacy (where the focus is on preventing companies from using Artificial Intelligence to guess your weaknesses, manipulate your decisions, and threaten democracy).

7.5. Synthesis of the Logical Chain between Inference, Law (Regulatory Gap), and Democracy

- Inference: profiling systems convert behavioral signals into inferred attributes and vulnerabilities, feeding recommendations and microtargeting at scale,

thereby generating epistemic and behavioral harms.

- Law: Data protection regimes focused on inputs tend to penalize failures in collection, governance, and security, but offer weaker instruments to govern the production and use of inferential outputs, reinforcing the input-output asymmetry (Saemann et al., 2022; Wachter & Mittelstadt, 2019; Veale & Borgesius, 2021).
- Democracy: when political messages and disinformation are tailored to inferred vulnerabilities, the fragmentation of the informational environment intensifies and the transparency of political persuasion is reduced, eroding the epistemic conditions of deliberation and the formation of free and informed will—in short, undermining democratic autonomy (Vosoughi, Roy, & Aral, 2018; Votta, Kruschinski, & Hove, 2024; Leerssen et al., 2023).

8. Conclusion

The core of the contemporary problem is not merely “personal data being collected,” but the ability to infer attributes and vulnerabilities to modulate behavior and attention. The economic scale of the targeted advertising ecosystem reinforces that inference is infrastructure, not the exception (Meta Platforms Inc, 2025; Alphabet Inc., 2026). Large-scale evidence indicates the structural advantages of diffusion for falsehoods in digital environments (Vosoughi, Roy, & Aral, 2018), while data on political advertising demonstrate the transnational expansion of microtargeting practices (Votta, Kruschinski, & Hove, 2024). In parallel, enforcement patterns suggest a focus on sanctions targeting traditional data cycle failures, with less traction on inferential harm (Saemann et al., 2022). Recent peer-reviewed scholarship confirms that the doctrinal architecture of both the GDPR and the DSA is structurally underequipped to address the epistemic and behavioral harms produced by inference-driven targeting (Wachter, Mittelstadt, & Russell, 2021; Veale & Borgesius, 2021; Leerssen et al., 2023).

Therefore, regulatory evolution needs to shift the center of gravity: from “data flows” to “knowledge power.” The right not to be inferred appears as a sector-specific legal instrument, a *lex specialis* operative in high-impact political and civic contexts, consistent with informational self-determination and, simultaneously, as a requirement to protect the minimum epistemic conditions of democratic autonomy in AI-mediated societies.

Acknowledgements

The author expresses sincere gratitude to Instituto Atlântico, where they serve as Data Protection Officer (DPO), for the institutional support and the academic freedom granted throughout this research. This autonomy was instrumental in developing the reflections of this paper regarding Artificial Intelligence, data regulation, inference governance, and systemic risks.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Alphabet Inc (2026). *Form 10-K 2024 (Annual Report)*.
- Barocas, S., & Nissenbaum, H. (2014). Big Data's End Run around Anonymity and Consent. In J. Lane, V. Stodden, S. Bender, & H Nissenbaum (Eds.), *Privacy, Big Data, and the Public Good* (pp. 44-75). Cambridge University Press.
<https://doi.org/10.1017/cbo9781107590205.004>
- Bioni, B. R. (2019). *Personal Data Protection: The Function and Limits of Consent*. Forense.
- De Miranda, P. R. (2022). *Right to Reasonable Inferences as a Normative Substrate in the Consolidation of Algorithmic Governance Mechanisms in Automated Decision Systems* (p. 129). Master's Thesis, Universidade Federal de Santa Maria.
- EDPB (2021). *Guidelines 8/2020 on the Targeting of Social Media Users (Version 2.0)*.
- European Union (2022). *Regulation (EU) 2022/2065 (Digital Services Act)*.
- Floridi, L. (2014). *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford University Press.
- Grand View Research (2026). *Data Broker Market Size and Share (Industry Report)*.
- Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law*. Edward Elgar Publishing. <https://doi.org/10.4337/9781849808774>
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Leerssen, P., Ausloos, J., Zarouali, B., Helberger, N., & De Vreese, C. (2023). Platform Transparency for the Public Interest: The Case for Public Audits of Social Media Platforms. *Internet Policy Review*, vol. 12.
- Lynskey, O. (2015). *The Foundations of EU Data Protection Law*. Oxford University Press.
- Meta Platforms Inc (2025). *Meta Reports Fourth Quarter and Full Year 2024 Results*.
- Pasquale, F. (2015). *The Black Box Society*. Harvard University Press.
<https://doi.org/10.4159/harvard.9780674736061>
- Rouvroy, A., & Pouillet, Y. (2009). The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy. In S. Gutwirth, Y. Pouillet, P. De Hert, C. de Terwangne, & S. Nouwt (Eds.), *Reinventing Data Protection?* (pp. 45-76). Springer. https://doi.org/10.1007/978-1-4020-9498-9_2
- Saemann, M., Theis, D., Urban, T., & Degeling, M. (2022). Investigating GDPR Fines in the Light of Data Flows. *Proceedings on Privacy Enhancing Technologies*, 2022, 314-331.
<https://doi.org/10.56553/popets-2022-0111>
- Solove, D. J. (2013). Privacy Self-Management and the Consent Dilemma. *Harvard Law Review*, 126, 1880-1903.
- Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press. <https://doi.org/10.1515/9781400884711>
- Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act—Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach. *Computer Law Review International*, 22, 97-112.
<https://doi.org/10.9785/cr-2021-220402>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The Spread of True and False News Online. *Science*, 359, 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Votta, F., Kruschinski, S., Hove, M., Helberger, N., Dobber, T., & De Vreese, C. (2024). Who Does(n't) Target You? Mapping the Worldwide Usage of Online Political Microtargeting. *Journal of Quantitative Description: Digital Media*, 4, 1-73.

<https://doi.org/10.51685/jqd.2024.010>

- Wachter, S., & Mittelstadt, B. (2019). A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 2019, 1-130.
- Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why Fairness Cannot Be Automated: Bridging the Gap between EU Non-Discrimination Law and AI. *Computer Law & Security Review*, 41, Article ID: 105567. <https://doi.org/10.1016/j.clsr.2021.105567>
- Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Council of Europe.
- WP29 (Article 29 Data Protection Working Party) (2018). *Guidelines on Automated Individual Decision-Making and Profiling (WP251rev.01)*.
- Zarouali, B., Dobber, T., & Schreuder, J. (2023). Personality and Susceptibility to Political Microtargeting: A Comparison between a Machine-Learning and Self-Report Approach. *Computers in Human Behavior*, 151, Article ID: 108024. <https://doi.org/10.1016/j.chb.2023.108024>
- Zarouali, B., Dobber, T., De Pauw, G., & de Vreese, C. (2022). Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media. *Communication Research*, 49, 1066-1091. <https://doi.org/10.1177/0093650220961965>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs.