# More Regulation, Please: Technical Privacy Protection Alone Can Only Partially Limit Democratic Harm from Microtargeting

**Position Statement**

Peter Schoo [1]

**Abstract:** This position statement identifies research questions and regulatory actions to mitigate the democratic risks of microtargeting, assuming the persistence of the data brokerage market. It argues that while a Privacy Enhancing Technology (PET) reduce privacy risks, PETs alone cannot address structural challenges like power asymmetries and unregulated data markets. Complementary regulatory measures are essential.

**Keywords:** GenAI, PETs, Microtargeting, Democracy

## 1 Democratic Discourses

GenAI has the potential to fundamentally reshape society, which embraces accelerating scientific discovery as well as democratising access to knowledge. However, this potential is increasingly at risk of being undermined by microtargeting. As for example discussed by Adrienne Fichter in *Smartphone-Democracy* (2017), personalised and manipulative content can distort democratic discourse, amplify polarisation, and erode individual autonomy [Fi17]. The book description repeatedly identifies microtargeting as one of the central mechanisms reshaping democratic discourse alongside fake news, bots, algorithmic polarisation, and messengerisation. In a further example [Zu18] with focus on political microtargeting a lack of data broker regulation and insufficient transparency is identified.

Against this backdrop, technical data protection approaches, and in particular PETs, are often proposed as a solution [Ag21][CMK24]. While they are an essential component of a modern data governance framework, their impact on democratic risks remains inherently limited.

[1] Independent Researcher, 82194 Gröbenzell, Germany,
peter.schoo@posteo.de, https://orcid.org/0009-0000-6147-2851

## 2 Intervention Points of Approach: Technical and Governance Perspectives on AI-Enabled Misinformation

The increasing efficiency and effectiveness of AI-enabled misinformation require interventions that reduce its societal impact. While no single measure can address this challenge comprehensively, computer science can make particularly strong contributions in two areas: PETs, which limit data-driven manipulation, and proposals for regulatory frameworks, whose design can be based on technical evidence and expertise.

1. *AI-enabled misinformation as a new-scale challenge*
   GenAI increases the speed, scale, personalisation, and effectiveness of misinformation campaigns. The challenge is therefore not only false content itself, but the ability to produce and distribute targeted influence operations at unprecedented efficiency [Pe26].
2. *From content detection to socio-technical resilience*
   Addressing AI-driven misinformation requires more than detecting individual pieces of content. Effective mitigation must consider the broader ecosystem of data, targeting, platforms, and amplification mechanisms.
3. *PETs as a technical contribution*
   PETs can reduce some enabling factors of manipulation by limiting excessive profiling, data exploitation, and privacy-invasive targeting. However, they cannot prevent the creation or dissemination of synthetic misinformation.
4. *Regulation as a governance contribution*
   Many risks arise from platform incentives, data markets, and accountability gaps. Regulation can establish transparency, responsibility, and risk management requirements, supported by technical expertise from computer science.
5. *Complementary measures and limitations*
   Other measures, such as media literacy and moderation, remain essential but face practical limitations in scalability and speed. Therefore, technical and regulatory interventions are important components of a broader resilience strategy.

This position statement thus focuses on the two intervention points PETs and regulation, while acknowledging that broader measures, such as digital literacy remain important complementary components of societal resilience, while content moderation is not seen as an option due to scale and speed of AI generated content. We argue how PETs and regulation can contribute to mitigating AI-enabled misinformation and identifies research questions at the intersection of privacy, data markets, platform governance, and democratic resilience.

## 3 The Role of PETs: Necessary but Not Sufficient

PETs, including differential privacy, federated learning, and homomorphic encryption, among others [In25], represent a mature class of technologies designed to minimise or prevent the re-identification of individuals in data processing. Other privacy-preserving approaches, such as VPNs, the Tor browser, encrypted communication services, and tracker-blocking tools, allow users to actively choose and control how their online activities are protected.

PETs primarily aim to prevent or significantly hinder the re-identification of individuals by employing techniques that obscure or eliminate direct and quasi-identifiers in datasets. Unlike user-controlled technologies such as VPNs or Tor, many PETs are implemented within data processing infrastructures and therefore protect privacy by design, independent of individual user decisions. PETs align closely with regulatory principles like data minimisation and "privacy by design" under the GDPR, ensuring that only the necessary data is processed in a way that inherently protects user identities.

PETs reduce the risk of exposing personal data, helping organisations comply with legal requirements while enabling data utility for analysis. However, they do not address the core mechanism of microtargeting, which are segmentation and behavioural influence based on inferred profiles, so that targeted manipulation remains possible. PETs restrict data processing but not the operational use of insights (patterns, preferences, or predictions) for behavioural influence.

To mitigate potential harm in advertising, politics, or social media, regulatory and ethical frameworks (e.g., bans on political microtargeting, transparency requirements) are needed alongside technical solutions.

#### The Invisible Foundation of GenAI is Data Quality

The long-term societal value of GenAI depends critically on the quality of its training data. A Large Language Model (LLM) inherits both the strengths and deficiencies of the datasets on which it is trained [Ch25].

A growing reliance on synthetic or unverified data introduces several risks:

- declining reliability of outputs due to hallucinations [ATN25],
- reduced economic value for organisations deploying GenAI [Ga24],
- increasing dependence on high-quality external data sources to stabilise performance [Ka25].

This dynamic creates a structural tension: while data protection aims to limit data availability, high-quality AI systems depend on robust, reliable datasets. This encompasses also misuse

of data that was originally given for a different purpose. The GDPR's purpose limitation principle requires that personal data be collected for specified, explicit, and legitimate purposes and not further processed in a manner incompatible with those purposes [Be24].

### Microtargeting: A Structural Threat to Democratic Systems

In parallel to data quality concerns, microtargeting poses a direct and systemic risk to democratic processes. By design, it enables the delivery of highly personalised content tailored to individual psychological profiles, often with persuasive or manipulative intent.

A central enabling factor is the largely unregulated ecosystem of data brokerage. Data brokers aggregate, infer, and trade personal data at scale, creating the informational foundation for targeted influence operations. Despite their systemic relevance, data brokers remain insufficiently regulated in practice [Ru23; SI25].

Microtargeting thus operates not merely as a technical feature, but as a power structure, i.e. it shifts informational asymmetries in favour of those controlling data and distribution channels [VMP25].

### PETs in the Training Phase of GenAI

When training a LLM with differential privacy, noise is added to the training process (e.g., gradients) to ensure the final model does not reveal specific details about individual training data points (re-identification), though this may slightly reduce the model's accuracy due to the interference of noise with its learning capacity.

Federated learning prevents centralised entities like tech corporations or data brokers from accessing raw data by training models locally on user devices and only aggregating updates, and it remains vulnerable to poisoning attacks or membership inference risks.

Additionally, using synthetic data—while not a PET in the strict sense—eliminates the need for real user data in training, thereby limiting the precision of microtargeting models.

Together, these approaches can enhance privacy in the training phase of GenAI, yet it introduces trade-offs between data utility and protection.

### PETs in the Inference Phase of LLMs

During the inference phase, differential privacy can be applied to queries by adding noise to model outputs, reducing the precision of personalised content and preventing direct

inferences about individual training data, such as shifting from hyper-personalised ads to broader segments like “interest in climate topics”.

Secure Multi-Party Computation allows multiple parties to collaboratively compute personalised recommendations without exposing their underlying data, though its high computational cost limits real-time microtargeting applications.

Homomorphic Encryption enables LLMs to process encrypted user data without decryption, ensuring raw data remains hidden, but its slow performance and resource demands restrict scalability for mass targeting.

#### Application Limits

All in all, PETs do not eliminate the core logic of microtargeting, which segments users based on inferred profiles, since PETs only limit data granularity and access. Data brokers or platforms may still combine aggregated PET-processed data with other sources to attempt re-identification through linkage attacks.

## 4 Structural Drivers of Influence

A narrow focus on technical privacy risks obscures the broader structural nature of the problem. Microtargeting persists even under reduced data granularity, particularly where platforms possess extensive first-party data reserves, such as Meta Platforms or Google.

Moreover, PET adoption is uneven and often requires active user engagement, limiting its effectiveness at population scale. As a result, existing power asymmetries between individuals and data-driven platforms remain largely intact.

In a multi-layered approach, to meaningfully reduce democratic risks, following changes are indicated:

#### Regulatory Intervention

- Restrictions or bans on political microtargeting [2]
- Mandatory transparency for political advertising
- Stronger enforcement of existing frameworks such as the Digital Services Act

[2] The case of the European Commission [Vo24] advertises a campaign on social networks promoting the Chat Control proposal shows how political actors can use personal data to try to influence specific voter groups rather than engaging in transparent public debate.

Regulate microtargeting now with bans, transparency obligations, and consistent enforcement of the Digital Services Act. Enforce, in particular, independent research into systemic risks in the digital environment and the accountability of platforms.

**Platform Governance**

- Reduction of manipulative recommender system dynamics
- Clear labelling of political and synthetic content

Reform platform governance and reduce manipulative algorithms and enforce clear labelling of political and synthetic content.

**Data Economy Reform**

- Specific regulatory regimes for data brokers as they create large-scale profiling and targeting infrastructures
- Secure data access frameworks for independent research and auditing

Reform the data economy to establish strict broker regulations and enable secure research access.

**Societal Capacity Building**

And last but not least

- Strengthening users' media literacy and awareness of targeting mechanisms and its risks

## 5 Research Questions and Call to Action

The analysis of this position statement highlights both the strengths and limitations of PETs that can substantially reduce privacy risks by limiting access to personal data, reducing re-identification risks, and supporting privacy-preserving data processing. Similarly, user-controlled technologies such as VPNs, Tor, encrypted communication services, and tracker-blocking tools allow individuals to reduce their digital exposure and limit data collection.

However, these technologies do not eliminate the core mechanisms that enable AI-driven microtargeting and influence operations. PETs controlled by data controllers primarily govern how data is processed, while user-controlled PETs primarily govern how data is

disclosed. Neither category directly addresses the economic incentives of data brokerage, the concentration of first-party data by large platforms, nor the operational use of insights for behavioural influence. As a result, PETs can reduce privacy risks while leaving democratic risks only partially mitigated.

This gap motivates following research questions focusing on the interaction between PETs, data markets, and democratic resilience.

> **RQ1** To what extent can user-controlled privacy technologies reduce exposure to profiling and microtargeting at population scale, and what barriers limit their adoption and effectiveness?

This question addresses whether empowering users with privacy technologies can meaningfully shift existing power asymmetries between individuals and data-driven platforms.

> **RQ2** To what extent can PETs deployed by data controllers reduce the feasibility and effectiveness of privacy-invasive profiling, microtargeting, and data brokerage while preserving legitimate data-driven innovation and services?

This question examines whether privacy-preserving processing can reshape business incentives behind large-scale profiling and behavioural targeting, beyond reducing re-identification risks. It operationalises this in distinguishable dimensions: the impact of PETs on targeting accuracy (maintaining or reducing profiling capabilities), broker profitability (changes in the value and monetisation of user data), and campaign effectiveness (the ability to achieve comparable advertising outcomes under privacy-preserving models).

> **RQ3** How can PETs and related accountability mechanisms be designed to make purpose limitation technically verifiable and auditable across complex data ecosystems involving platforms, AI providers, and data brokers?

While the GDPR restricts the use of personal data to specified purposes, demonstrating unlawful secondary use remains difficult in practice. This question investigates whether techniques such as data provenance, usage control, cryptographic auditing, and machine-readable consent mechanisms can provide verifiable evidence that personal data is processed only for the purposes originally authorised by the data subject.

> **Regulatory Action** should move beyond requiring consent and transparency alone. Organisations that process personal data at scale should be required to provide technically verifiable evidence of purpose compliance through auditable data provenance, usage logs, and independent verification mechanisms. Similar to financial auditing, privacy compliance should become demonstrable rather than merely declarative.

From a computer science perspective, future work should provide evidence-based metrics and technical mechanisms that enable regulators to assess the effectiveness of such measures and their impact on democratic resilience.

## 6 Conclusion

This position statement highlights the need for both technical solutions and regulatory changes to address the democratic risks of microtargeting. It identifies key research questions for computer scientists to explore the interplay between PETs, data markets, and democratic resilience, emphasising that technical measures alone are insufficient without robust governance and therefore stronger regulation that is indicated, as PETs alone can't do.

## Declaration of AI Assistance

GenAI assisted in the preparation of this position statement for both fact-finding research and linguistic drafting.

## References


[Ag21] Agrawal, N. et al.: Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21: CHI Conference on Human Factors in Computing Systems. ACM, Yokohama Japan, pp. 1–13, 2021.

[ATN25] Anh-Hoang, D.; Tran, V.; Nguyen, L.-M.: Survey and Analysis of Hallucinations in Large Language Models: Attribution to Prompting Strategies or Model Behavior. Frontiers in Artificial Intelligence 8, p. 1622292, 2025.

[Be24] Becker, R. et al.: Purpose Definition as a Crucial Step for Determining the Legal Basis under the GDPR: Implications for Scientific Research. Journal of Law and the Biosciences 11 (1), lsae001, 2024.

[Ch25] Chen, Z. et al.: Revisiting Scaling Laws for Language Models: The Role of Data Quality and Training Strategies. In: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, Vienna, Austria, pp. 23881–23899, 2025.

[CMK24] Calvi, A.; Malgieri, G.; Kotzinos, D.: The Unfair Side of Privacy Enhancing Technologies: Addressing the Trade-Offs between PETs and Fairness. In: The 2024 ACM Conference on Fairness, Accountability, and Transparency. FAccT '24: The 2024 ACM Conference on Fairness, Accountability, and Transparency. ACM, Rio de Janeiro Brazil, pp. 2047–2059, 2024.

[Fi17] Fichter, A., ed.: Smartphone-Demokratie: #Fake News, #Facebook, #Bots, #Populismus, #Weibo, #Civic Tech. NZZ Libro, Neue Zürcher Zeitung, Zürich, 2017.

[Ga24] Gartner, Inc.: Gartner Predicts 30% of Generative AI Projects Will Be Abandoned After Proof of Concept By End of 2025, Gartner, 2024, https://www.gartner.com/en/newsroom/press-releases/2024-07-29-gartner-predicts-30-percent-of-generative-ai-projects-will-be-abandoned-after-proof-of-concept-by-end-of-2025, accessed: 05/05/2026.

[In25] Information Technology and Innovation Foundation: ITIF Technology Explainer: What Are Privacy Enhancing Technologies?, 2025, https://itif.org/publications/2025/09/02/itif-technology-explainer-privacy-enhancing-technologies/, accessed: 05/05/2026.

[Ka25] Kang, F. et al.: Demystifying Synthetic Data in LLM Pre-training: A Systematic Study of Scaling Laws, Benefits, and Pitfalls. In: Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Suzhou, China, pp. 10750–10769, 2025.

[Pe26] Peters, J.: Die Propaganda-KI der AfD, correctiv.org, 2026, https://correctiv.org/aktuelles/geschaefte-der-afd/2026/07/02/die-propaganda-ki-der-afd/, accessed: 07/23/2026.

[Ru23] Ruschmeier, H.: Data Brokers and European Digital Legislation. European Data Protection Law Review 9 (1), pp. 27–38, 2023.

[SI25] Sebastian Meineck; Ingo Dachwitz: Targeting the EU, netzpolitik.org, 2025, https://netzpolitik.org/2025/databroker-files-targeting-the-eu/, accessed: 05/05/2026.

[VMP25] Van Gils, F.; Müller, W.; Prüfer, J.: Microtargeting, Voters' Unawareness, and Democracy. The Journal of Law, Economics, and Organization 41 (2), pp. 634–653, 2025.

[Vo24] Volker Briegleb: Advertising for Chat Control: EU Data Protection Watchdog Rebukes Commission | Heise Online, 2024, https://www.heise.de/en/news/Advertising-for-chat-control-EU-data-protection-watchdog-rebukes-Commission-10199555.html, accessed: 07/20/2026.

[Zu18] Zuiderveen Borgesius, F. J. et al.: Online Political Microtargeting: Promises and Threats for Democracy. Utrecht Law Review 14 (1), p. 82, 2018.