GUIDANCE IN ANONYMIZATION: WHEN AMBIGUITY MEETS PRIVACY-WASHING

Szilvia Lestyán
szilvia.lestyan@inria.fr

William Letrone
william.letrone@univ-nantes.fr

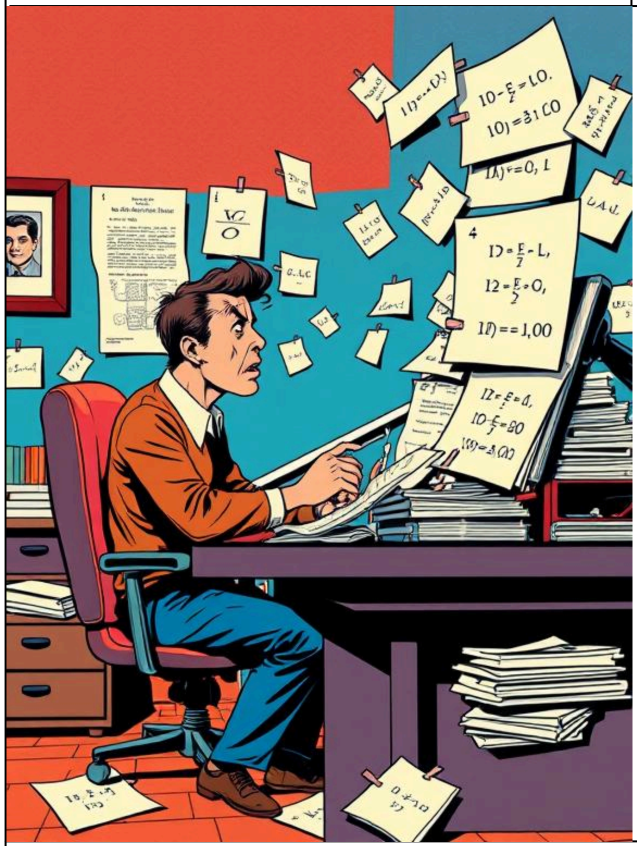
Ludovica Robustelli
ludovica.robustelli@univ-nantes.fr

INTRODUCTION

ANONYMIZATION'S EFFECTIVENESS REMAINS INCONSISTENT DUE TO OUTDATED TECHNIQUES, UNCLEAR REGULATIONS, AND PRACTICAL MISAPPLICATIONS. THIS PAPER EXPLORES PRIVACY-WASHING FROM THE ANONYMIZATION PERSPECTIVE, WHERE ORGANIZATIONS CLAIM ADEQUATE SAFEGUARDS WHILE FAILING TO PROVIDE MEANINGFUL PRIVACY

OBJECTIVES

- Highlight regulatory ambiguities, outdated techniques, and educational gaps that contribute to privacy-washing.
- By analyzing guidelines, case studies, technical documentations, and legal frameworks, we point out the underlying causes of privacy-washing.
- We propose solutions to bridge the gap between regulation and practice.
- Our goal is to ensure that applied anonymization techniques align with modern privacy threats and technological advancements, offering better protections for both individuals and organizations.



KEY CHALLANGES

REGULATORY AMBIGUITY

➢ Inconsistent interpretations of anonymization under GDPR across EU Member States create uncertainty. Some authorities adopt a strict approach, while others allow for more flexible compliance, leaving businesses unsure of best practices.

➢ OUTDATED TECHNIQUES

Many organizations continue to rely on traditional methods such as k-anonymity and l-diversity, despite well-documented weaknesses. These approaches fail to account for modern re-identification attacks, which leverage auxiliary data sources and machine learning techniques.

LACK OF PRACTICAL GUIDANCE

➢ Most available guidance is either high-level (business oriented) or very technical (scientific literature), making it difficult for practitioners to apply anonymization effectively. Engineers and data handlers lack access to practical, step-by-step instructions on implementing privacy-preserving techniques.

REGULATIONS AND GUIDELINES

EU: The GDPR interpretations differ among DPAs (CNIL vs. ICO or DPC)
EDPB still refers to **WP29**, despite many critiques. [4]
Guidelines on **AI** privacy: "*personal data cannot be inferred*" - no technical guidelines exists.
Guidelines on **pseudonymisation**: no mention of anonymization, source of common confusion [1]
Misleading or wrong examples are published by DPAs. [5]
➢ Need for clearer guidance. [2]

Global Perspective: Outside the EU, regulatory frameworks differ significantly. The U.S. relies on HIPAA and CCPA, which define de-identification differently from GDPR. Japan's APPI and Brazil's LGPD offer yet another interpretation. These differences create compliance challenges for multinational organizations handling personal data.
➢ *Incompatibility of legal regimes as main challenge to cross-border data flows*

CASE STUDIES

Ethical codes: having organizational measures or ethical codes is not enough to satisfy data protection principles. (Garant vs INPS)

Mislabeled Data: Several organizations have been found mislabeling pseudonymized data as fully anonymized, leading to regulatory penalties and loss of consumer trust. (CNIL vs Doctissimo)

Legal Disputes: Courts have ruled against companies that claimed their anonymization practices complied with GDPR but failed to prevent re-identification. (Garant vs Camedi)

Compliance Struggles: Businesses struggle to navigate complex anonymization guidelines, leading to inconsistent implementation and privacy risks. (CNIL vs Cegedim)

RECOMMENDATIONS

Improved Education: develop a comprehensive data privacy curriculum that could be promoted and distributed by data protection authorities either in a form of offered educational programs tailored to diverse audiences.
(1) A clear explanation of privacy threats,
(2) An overview privacy definitions,
(3) Application and evaluation,
(4) Best practices,
(5) Case studies,
(6) Hands-on learning exercises etc.

Better Auditing Tools: Privacy risk assessment frameworks should be integrated into compliance processes to help organizations evaluate whether their anonymization techniques are effective. [3]

[1] AEPD EDPB. 2021. 10 misunderstandings related to anonymisation.
[2] European Commission EC. 2024. Communication from the Commission to the European Parliement and the CouncilL - Second Report on the application of the General Data Protection egulation,COM/2024/357.
[3] Andrea Gadotti, Luc Rocher, Florimond Houssiau, Ana-Maria Crețu, and Yves-Alexandre De Montjoye. 2024. Anonymization: The imperfect science of using data while preserving privacy. Science Advances 10, 29 (2024), eadn7053.
[4] Sophie Stalla-Bourdillon and Alison Knight. 2016. Anonymous data v. personal data-false debate: an EU perspective on anonymization, pseudonymization and personal data. Wis. Int'l LJ 34 (2016), 284.
[5] ICO. 2023. Privacy-enhancing technologies (PETs). - Case study: differentially private mixed noise in financial services

*Images were generated by deepai