

United Nations AI For Good Summit • ITU Kaleidoscope Academic Conference 2026

PROTECTING SYNTHETIC TELECOMMUNICATIONS DATA: A LEGAL REVIEW AND PROPOSED REFORMS

Alex Austria

*Master of Laws in Technology, Media, and Telecommunications
Queen Mary University of London*



WHAT IS SYNTHETIC DATA?

*Synthetic data refers to data that we create to **mimic** the properties and patterns of real-world data. We **maintain** the statistical properties, relationships and distributions of that original data, often with some **corrections or modifications** that make it especially beneficial in the development and tuning of machine learning models.*

– From “AI Insights; Synthetic Data”,
The UK Government Digital Service, updated 04 November 2025





SYNTHETIC TELECOMMUNICATIONS DATA AND RE-IDENTIFICATION

- Especially vulnerable to it
- AI makes it easier to re-identify



SYNTHETIC DATA'S GOVERNANCE PROBLEM

Can the data controller demonstrate that the re-identification risk has been:

↳ reduced to a a materially low level?

↳ technically tested?

↳ documented through appropriate evidence?

↳ periodically reassessed?





THE TRIPARTITE TEST OF IDENTIFIABILITY



SINGLING OUT



The ability to isolate records relating to an individual within a dataset

LINKABILITY



The ability to connect records concerning the same person within one dataset or across datasets

INFERENCE



The ability to deduce, with significant probability, an attribute from other attributes



MAIN POINT:

Synthetic telecommunications data should be treated as no longer relating to an identifiable natural person **only where the residual risk of singling out, linkage, or inference**, assessed against state-of-the-art means reasonably foreseeable in the sectoral threat model, is **materially lower than the risk tolerated** in equivalent pseudonymous processing and is supported by documented PET application, testing, and governance



PROPOSED LEGAL REFORMS



TIER 1: **Interpretative** **guidance under** **Art. 70, EU GDPR**



TIER 2: **Binding duties in** **operative** **instruments**



Art. 89, EU GDPR

Tripartite test is the operative mechanism; specify how Arts. 26 and 4(1), EU GDPR should be applied where synthetic or de-identified telecommunications data is used

PETs should be incorporated in binding instruments; Art. 10 of EU AI Act; EHDS as a sectoral analogy

The tripartite test sets the threshold, and Article 89 governs the processing that remains above it

A vertical image on the left side of the slide showing two telecommunications towers against a blue sky with white clouds. The towers are covered in various antennas and equipment. A network of glowing green lines connects different points, suggesting a data or communication network.

POLICY RECOMMENDATIONS FOR REGULATORS AND TELECOM OPERATORS

-  Develop sector-specific anonymization and synthetic data guidance
-  Conduct synthetic DPIAs
-  Secure access environments and provide more proportionate safeguards
-  Validate and combine PETs
-  Periodic assessment of standards



CONCLUSION

Synthetic data can be a powerful privacy-enhancing tool, but only when its risks are tested, its safeguards are documented, and its status is periodically re-evaluated.

By codifying tripartite identifiability testing, telecommunications data governance can preserve the utility of synthetic data while altogether preserving subscriber trust.



THANK YOU!

m.c.austria@hss25.qmul.ac.uk