The missing link between AI innovation and data protection?

Balancing innovation, compliance, and trust in the age of AI-driven data processing

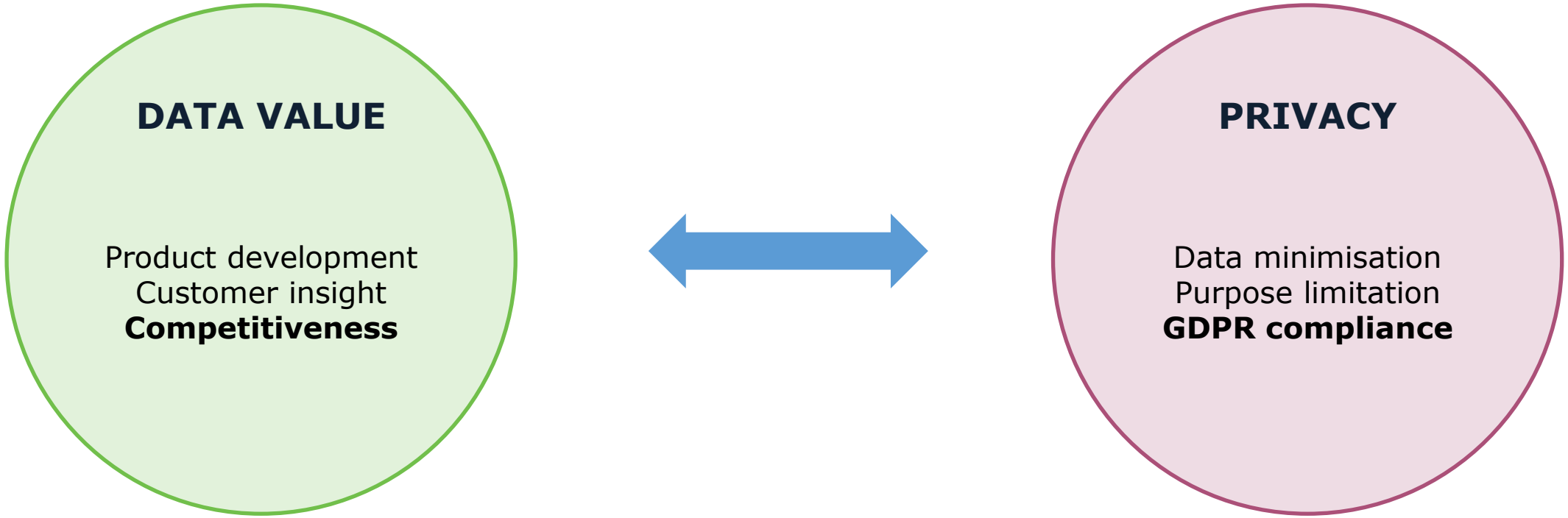


Dr. Eleni Alexiou, CIPP/E, AIGP
Dr. Tina Gausling, LL.M. (Columbia), CIPP/E

IAPP | European Privacy KnowledgeNet 2026

The business tension

More data creates value. More personal data creates exposure.



Synthetic data aims to preserve useful structure while breaking the link to real individuals.

What synthetic data actually means

Method, output and legal status must be kept separate.

SYNTHETIC DATA

Artificial records generated from learned distributions or simulation rules.

Preserves selected patterns and correlations.

Describes provenance and method.

ANONYMOUS DATA

No identified or identifiable natural person from the applicable perspective.

Requires identification risk to be insignificant in practice.

A legal and risk-based conclusion.

PSEUDONYMOUS DATA

Identifiers are replaced or separated, but identification remains possible with additional information.

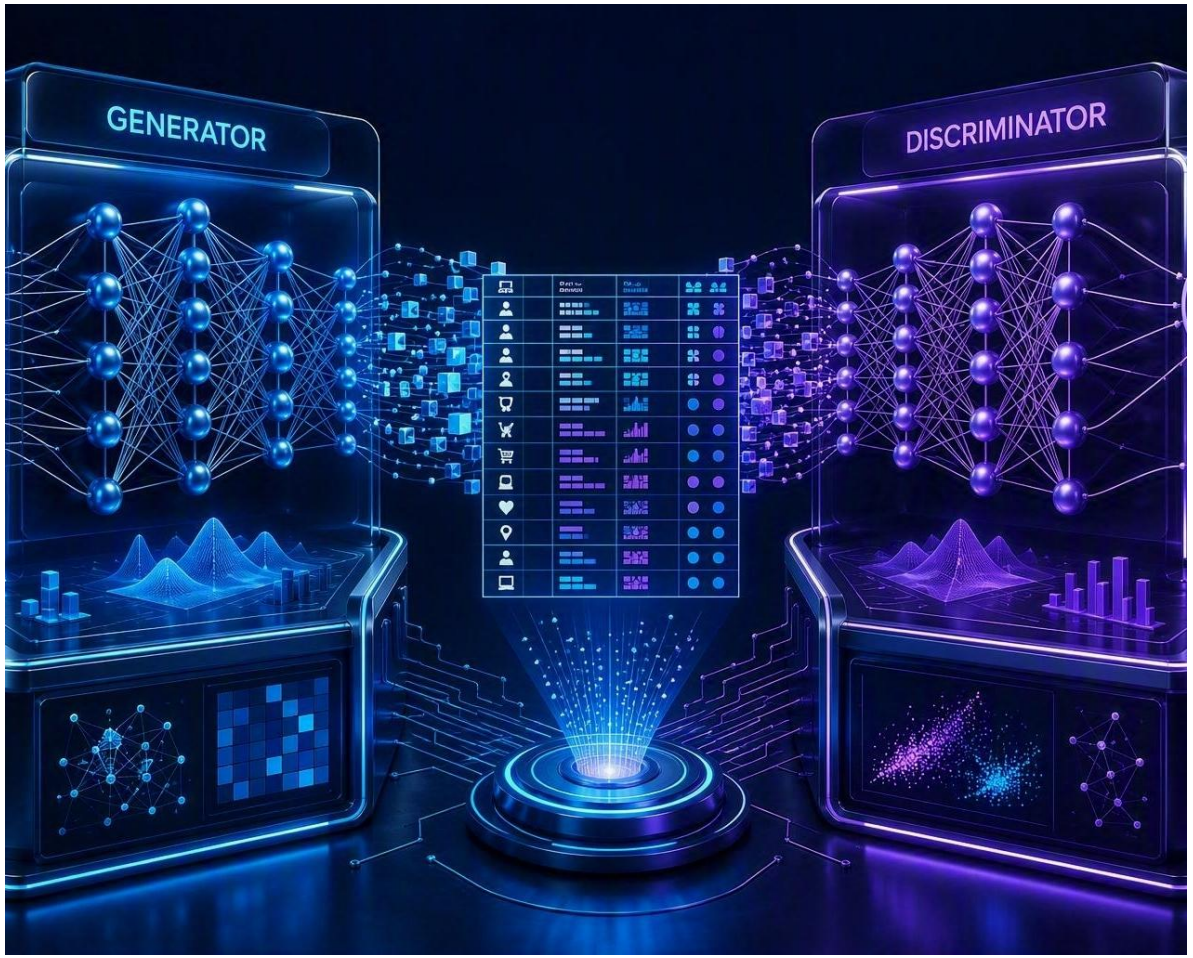
GDPR continues to apply.

A safeguard, not an exit.

GENERATION ≠ ANONYMISATION ≠ GUARANTEED ANONYMITY

How generation works: the GAN example

Two neural networks improve each other through an adversarial learning loop.



1. **Learn** — The generator learns patterns in the source data.
2. **Challenge** — The discriminator tests whether outputs look real.
3. **Iterate** — Feedback updates model parameters.
4. **Generate** — New records preserve patterns, not identities.

The EDPB test for anonymity

Can individuals be distinguished and treated differently with sufficient confidence?

01

NO RECORD ISOLATION

No unique combination of attributes relating to one individual.

02

NO LINKAGE

No effective connection to other information about the same individual.

03

NO INFERENCE

No specific and meaningful inference about an identified or identifiable individual.

Pass all three: anonymity is supported. Fail one: continue the analysis.

Protecting against membership inference

Reduce how much a model can reveal about whether one person was in the source data.

MINIMISE UNIQUENESS

Generalise rare attributes, cap outliers and avoid tiny cohorts.

CONTROL LEARNING

Regularise and stop before memorisation.

ADD PRIVACY

Apply differential privacy where appropriate and document the privacy budget.

Target state: the model behaves nearly the same whether one person is included or not.

Anonymity is perspective-specific

Ask first: for whom is the output intended to be anonymous?

ANONYMISING CONTROLLER

Source data, additional information and available means may make identification possible.

PROCESSOR

The controller’s perspective applies. Outsourcing does not change the data’s personal nature.

INDEPENDENT RECIPIENT

Its access, additional information and means reasonably likely to be used determine the result.

UNAUTHORISED ACTOR

Relevant where unlawful access or identification is reasonably likely in practice.

Two ways to apply the EDPB Framework

CONTEXTUAL APPROACH

Assess the relevant entity’s actual capabilities

SIMPLIFIED APPROACH

Disregard differences between entities for a more conservative assessment

The same dataset may be anonymous for one entity and personal for another.

Separate the two legal stages

The output may be anonymous, while the generation process remains fully regulated.

STAGE 1 | SYNTHESIS PROCESS

Personal source data is processed to generate the synthetic data set:

- requires an Article 6 GDPR legal basis and
- where relevant, an Article 9(2) GDPR condition
- any further purpose to be assessed under Article 6(4) GDPR compatibility test
- data minimization, security, transparency and accountability apply throughout the process



STAGE 2 | SYNTHETIC OUTPUT

Assess anonymity per relevant entity:

- Apply the three EDPB criteria: No record isolation, no linkage and no inference
- Test relevant identification techniques under realistic conditions
- Consider additional information and the means reasonably likely to be used by the relevant entity
- Periodically reassess the likelihood of re-identification.

The generation process remains regulated even if the validated output qualifies as anonymous.

Stage 1 | Legal basis for the synthetis process

The applicable route depends on whether the synthetis pursues the same or a further purpose

1 **SAME PURPOSE?**

- If synthesis serves the same purpose, the original legal basis may continue to apply.

2 **FURTHER PURPOSE?**

- Run the Article 6(4) compatibility test considering the link, context, data type, possible consequences and safeguards.

3 **ARTICLE 6(1)(f)?**

- Define a concrete legitimate interest, establish necessity, assess less intrusive alternatives and balance the effects on individuals.

4 **SPECIAL DATA?**

- A separate Article 9(2) condition is required where special-category data is processed.

IMMEDIATE ANONYMISATION AND DELETION OF ORIGINAL DATA MAY SUPPORT THE LEGITIMATE-INTEREST BALANCING TEST

(EDSA 02/2026 GUIDELINES, PARA. 39)

Stage 2 | Putting the anonymity assessment in practice

A defensible anonymity claim combines engineering, context and governance.

MAP

Map the given data and additional information that may aid re-identification.

SCOPE

Identify the relevant entities and means reasonably likely to be used from their perspective.

TEST

Test No Record Isolation, No Linkage and No Inference using relevant re-identification techniques.

EVALUATE

Assess whether identification results are sufficiently precise and reliable to distinguish an individual.

DOCUMENT

Document the anonymization process and testing of the output.

RE-TEST

Periodically reassess the likelihood of re-identification.

Anonymity conclusion should be supported by a documented technical assessment

Customer analytics under the AI Act

AI Act classification depends on the system’s intended purpose, not on whether its data is synthetic.

TYPICAL ANALYTICS

- Segment customers based on statistically relevant patterns
- Predict churn or the likelihood of tariff changes
- Stimulate purchasing decisions, click paths and customer journeys
- Test pricing strategies and cross-selling algorithms.

Usually not high-risk under the AI Act by itself.

POTENTIAL HIGH-RISK USE

AI system intended to evaluate the creditworthiness of natural persons or establish their credit score (Annex III no. 5(b) AIA)

If high-risk classification applies:

Article 10 AIA requires data governance, relevance, representativeness, sufficient completeness and data that is as error-free as possible.

GDPR: Personal or anonymous?

AI Act: Intended purpose and risk class?

If anonymity fails

The consequences may extend beyond the dataset to downstream processing, the model and its deployment.

GDPR CONTINUES

The synthetic dataset remains personal data, and its downstream use continues to be subject to the GDPR

ENFORCEMENT

- Supervisory authorities may impose fines, restrict further processing or order the deletion of unlawfully processed data.
- If personal information has entered a model, remediation may extend beyond the underlying dataset.

MODEL REMEDIATION

- Depending on the circumstances, remediation may include retraining, unlearning, removal of affected training data or additional privacy controls.
- Functional restrictions or deactivation may be necessary where risk cannot otherwise be addressed.

AI ACT OVERLAY

- If a high-risk AI system no longer complies with the applicable Article 10 data-governance requirements, the provider may need to take corrective action under Article 20 AI Act.
- Corrective action may affect the dataset, the model or the conditions under which the system is deployed.

Validate anonymity before release and model integration.

An integrated governance framework for synthetic data

jApply legal, technical and organizational controls throughout the lifecycle

1

USE CASE

Define the business objective, intended decision, relevant users and required analytical value

2

SOURCE

Confirm lawfulness, original purpose, data minimisation and source-data quality

3

SYNTHESISE

Select and document the method, parameters and technical privacy safeguards.

4

VALIDATE

Test re-identification risks, analytical utility, data quality, bias and potential distortions.

5

RELEASE

Confirm that the legal and technical assessment supports anonymity for the intended release context; access rights, permitted uses and onward sharing.

6

MONITOR

Monitor changes in data, linkage opportunities, recipients, incidents and available re-identification techniques.

Clear responsibilities • Internal Controls • Continuous quality assurance • Documented decisions

THREE TAKEAWAYS

01 Synthetic data can enable customer analytics and AI development while reducing reliance on real personal data.

02 The synthesis process requires a legal basis; the output leaves the GDPR only where re-identification is sufficiently unlikely from the relevant perspective.

03 For high-risk AI systems, synthetic data must meet the applicable data-governance and quality requirements for training, validation and testing datasets.

Synthetic data can enable privacy-conscious AI development and bridge innovation and data protection, but only where anonymity is demonstrated rather than assumed.

Questions & discussion

In Kooperation mit:
Bitkom
BvD
davit im DAV
eco
VAUNET

ZD
ZEITSCHRIFT FÜR DATENSCHUTZRECHT

Herausgeber: RA Prof. Dr. Jochen Schneider · Prof. Dr. Thomas Hoeren · Prof. Dr. Martin Selmayr · RA Dr. Axel Spies · RA Tim Wybitul

Newsdienst ZD Aktuell
Inklusive Online Newsdienst

AUS DEM INHALT

	121	THILO WEICHERT **Datenschutz bei PayPal**
	123	BARBARA SCHMITZ / JOCHEN SCHNEIDER / AXEL SPIES **Zwischen Grundrechtsschutz und Datenpraxis – Braucht die DS-GVO eine Reform?**
Re-Identifizierbarkeit	126	ELENI ALEXIOU / TINA GAUSLING **Synthetische Daten – Schlüssel zu datengetriebenen Unternehmensstrategien?**
Datennutzungsvertrag	132	MARTIN SCHEURER / BARBARA SCHMITZ **Datenschutzkonforme Vertragslösung: „Pay or Conclude"?**
Grenzüberschreitender Datentransfer	140	CONSTANTIN FRANK-FAHLE / MARCEL TROST **Das Datenschutzrecht der Vereinigten Arabischen Emirate**
Scoring	147	AXEL ANDERL / NINO TLAPAK / CORINA KRUESZ **Automatisierte Bonitätsprüfung**
Verantwortlichkeit	150	EuGH: **Haftung von Online-Marktplatz-Betreibern** mAnm HENSE
Direktmarketing	161	EuGH: **Newsletter-Werbung ohne Einwilligungserfordernis** mAnm ECKHARDT
Bodycam-Einsatz	175	BayVGH: **Einschreiten der Datenschutzaufsichtsbehörde** mAnm THIEL

www.zd-beck.de
linkedin.com/showcase/zeitschriftzd
C.H.BECK
Seiten 121–180
16. Jahrgang 2. März 2026
3/2026

You presenters today



Dr. Eleni Alexiou CIPP, AIGP

Head of Data Privacy Law
ADAC e.V.

Eleni.Alexiou@adac.de



Dr. Tina Gausling LL.M. (Columbia), CIPP/E

RAin, Fachanwältin IT-Recht
Simmons & Simmons

tina.gausling@simmons-simmons.com