

Synthetic data: implications for healthcare and data law

Elisabetta Biasin

KU Leuven Centre for IT & IP Law (CiTiP) – imec
Stanford Law School TTLF Fellow

Computers, Privacy & Data Protection (CPDP)

25/05/2023 8:45-10:00

Brussels

Contents

Introduction

Legal Implications of Synthetic Data

Synthetic Data in the EU Data Spaces: the EHDS

Conclusions and further discussions



Introduction



Synthetic data

(Rubin, 1993)

OECD:

'approach to confidentiality'

Data Governance Act

*'Privacy preserving methods
that could contribute to a more
privacy-friendly processing of
data'*

Synthetic data

EDPS (web):

*The concept of synthetic data generation is to take
an original data source (dataset) and create new
artificial data with similar statistical properties
from it.*

GDPR

How are Synthetic Data Used? (Healthcare)



Use: (Examples)

- Simulation and Prediction Research
- Hypothesis, Methods, and Algorithm Testing
- Epidemiological Study/Public Health Research
- Education and Training
- Public Release of Datasets
- Linking Data
- Health IT Development and Testing

(Gonzales, Guruswamy, Smith, 2023)

Use of Synthetic Data: Healthcare (in silico/AI)



Legal implications



Legal implications of synthetic data

DATA LAW

DATA PROTECTION

Pseudonymisation and anonymisation

- Personal or non-personal data?

(Fontanillo Lopez & Elbi, 2022a; 2022b)

Identifiability and re-identification

- See WP29 Guidance + Case law

SRB v EDPS

Data Minimisation & Purpose Limitation

- Further processing of personal data (+ data sharing)

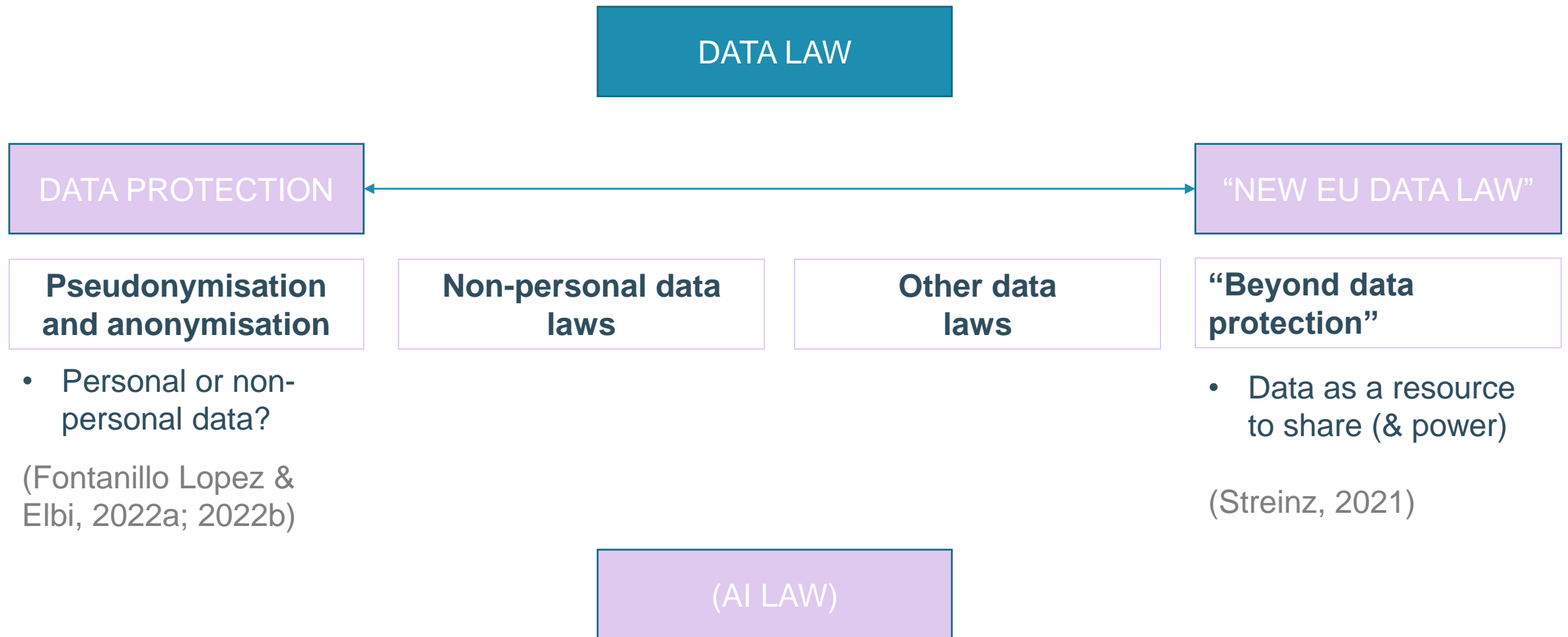
Accuracy and Data Quality

- To remedy paucity of data (>< nb harms)

(see also Gal & Lynskey, 2023;)

**Synthetic data
have implications that go beyond data *protection***

Legal implications of synthetic data



**Synthetic data
have implications that go beyond *data laws***

Other implications: healthcare (laws)

HEALTH LAW

MEDICAL DEVICE REGULATIONS

Qualification of synthetic data

- Should they and their results be considered medical devices?
- Model credibility/trust

PHARMACEUTICAL LAW

Clinical Trials reform

- Legal and regulatory pathways (eg *in silico* trials);
- Model credibility/trust

(Biasin, 2023; Viceconti et al 2020)

Uriel on Unsplash

Other implications: healthcare (ethics)

ETHICAL ASPECTS

BIOMEDICAL ETHICS PRINCIPLES

Principle of justice

- Synthetic data to address underrepresented populations (debated point)

Non-maleficence (do no harm) / Hy.Oath

- Possible harm (from data inaccuracies; data misuse from scientists)

(under analysis in the In Silico World project)

Uriel on Unsplash

Synthetic Data in the EU Data Spaces: the EHDS



**Question:
will synthetic data facilitate health data sharing in
the EHDS?**



GDPR

DATA
GOVERNANCE
ACT

EUROPEAN
HEALTH DATA
SPACE

(Slokenberga, 2022;
Ylmaz & Shabani
2022)

EHDS: Secondary Use of Health Data



Rules and mechanisms supporting the secondary use of electronic health data

Personal and non-personal

Synthetic data

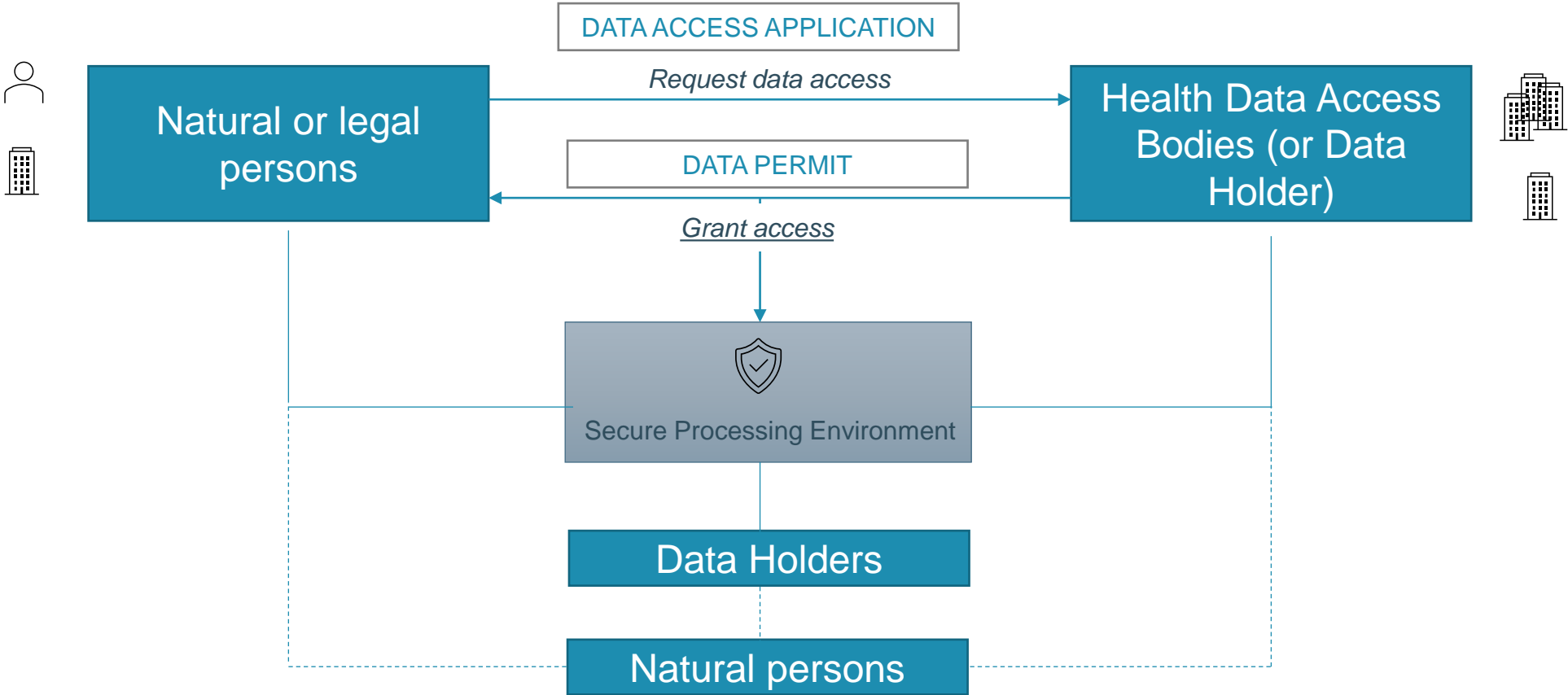
Proposal for a
REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
on the European Health Data Space

(Text with EEA relevance)

{SEC(2022) 196 final} - {SWD(2022) 130 final} - {SWD(2022) 131 final} -
{SWD(2022) 132 final}

1. **Scientific research about health or care research**
2. **Development of innovation activities** for products or services contributing to public health or social security, quality and safety of healthcare, medicinal products or medical devices
3. **Training, testing and evaluating of algorithms,**
4. Providing personalised healthcare
5. Public interest in the area of public health
6. Task of public sector body
7. Official statistics
8. Education or teaching activities

EHDS: Secure Processing Environment



(Biasin, 2023, ISW D9.2)

EHDS: Synthetic Data in Article 44

Article 44: Data minimization and purpose limitation



Health data access bodies shall provide the electronic health data in an **anonymized format**, where the purpose of processing can be achieved with such data.



When the purpose of the processing cannot be achieved with anonymized data, health data access bodies shall provide access to electronic health data in **pseudonymized format**



Prohibition for data users to re-identify the electronic health data provided in pseudonymized format.



Justification needed when pseudonymized data are requested. The data request shall additionally include:

- description of compliance with GDPR legal bases;
- information of ethical aspects of the processing

**Synthetic data
will play an increasing role in the health data space**

Conclusion



Conclusions

1

Legal Implications: implications in data protection law and beyond (example: healthcare). There are also **ethical and**

- **Doctrine:** ECJ ruling effects on synthetic data; aspects of synthetic data beyond DP
- **Legislator: EHDS** Article 44: extend guarantees to anonymised data
- **Regulators:** data protection (but also medical device and medicinal products guidance) should tackle them

Recommendations

Guilherme Stecanella on Unsplash

Conclusions

1

Legal Implications: implications in data protection law and beyond (example: healthcare). There are also **ethical and societal implications**.

2

Some recommendations

3

Conclusion / further discussion in the panel: In health data spaces – given the EHDS (and given the most recent ECJ developments) – synthetic will likely to play an increasing role



Guilherme Stecanella on Unsplash

Thank you!

KU LEUVEN

CiTiP

CENTRE FOR IT & IP LAW

PI: Prof. dr. Anton Vedder
KU Leuven Centre for IT IP Law

Elisabetta Biasin
Doctoral Researcher KU Leuven Centre for IT IP Law
Stanford Law School TTLF Fellow
EMA External Collaborating Expert on Data Protection of
Big Data and Real-World Data

Elisabetta.biasin@kuleuven.be

@bisilisib

+32 16 37 77 73

References

- Biasin E, In-depth analysis of legal and ethical requirements (2023)
- Fontanillo López CA and Elbi A, 'On Synthetic Data: A Brief Introduction for Data Protection Law Dummies' (2022)
- Fontanillo López CA and Abdullah E, 'On the Legal Nature of Synthetic Data' (2022)
- Gal M and Lynskey O, 'Synthetic Data: Legal Implications of the Data-Generation Revolution' [2023] SSRN Electronic Journal <<https://www.ssrn.com/abstract=4414385>> accessed 26 May 2023
- Gonzales A, Guruswamy G and Smith SR, 'Synthetic Data in Health Care: A Narrative Review' (2023) 2 PLOS Digital Health e0000082
- Streinz T, The Evolution of European Data Law In: The Evolution of EU Law Edited by: Paul Craig and Gráinne de Búrca, Oxford University Press. Paul Craig and Gráinne de Búrca 2021. DOI: 10.1093/oso/9780192846556.003.0029