Data Anonymization in Theory and Practice

Prof. Dr. Fabian Prasser

Medical Informatics Group
Berlin Institute of Health @
Charité – Universitätsmedizin Berlin

Berlin Institute of Health
Charité & MDC

Aus Forschung wird Gesundheit

Berlin Institute of Health
Charité & MDC

Outline

1. Background
2. Threats and protection methods
3. Anonymization of analysis results
4. ARX Data Anonymization Tool
5. Real-world examples

1. Background

Motivation

- Data sharing: Big data approaches in medical research
  - Precision medicine: high case numbers, detailed characterizations
  - Real-world evidence: secondary use, e.g. of routine clinical data for research
  - Collaborative research, e.g. data sharing across institutional boundaries
- Open science: Initiatives to improve the transparency, reproducibility and reusability of research results and research data
  - NIH Statement on Sharing Research Data, Notice NOT-OD-03-032; 2003.
  - NIH Genomic Data Sharing Policy, Notice NOT-OD-14-124; 2014.
  - EMA Policy 0070 on Publication of Clinical Data for Medicinal Products for Human Use; 2014.
- Data protection requirements

Background: Terminology and principles in the GDPR

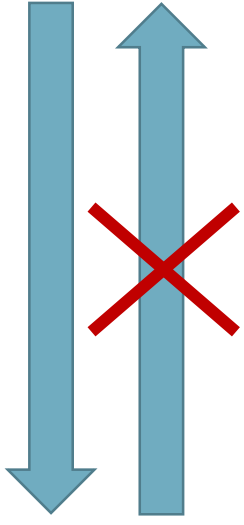
- **Terminology used in the regulation:** personal data, identified or identifiable person, anonymous data, pseudonymisation

- **Terminology *not used* in the regulation:** anonymisation, anonymised, pseudonymised, de-identification, de-identified, coded etc.

- **Principles:** lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity, confidentiality, accountability

Background: Anonymous data according to the GDPR

Personal data



Anonymous data

GDPR, Recital 26:

„The principles of data protection should **apply to any information concerning an identified or identifiable natural person** [...]“

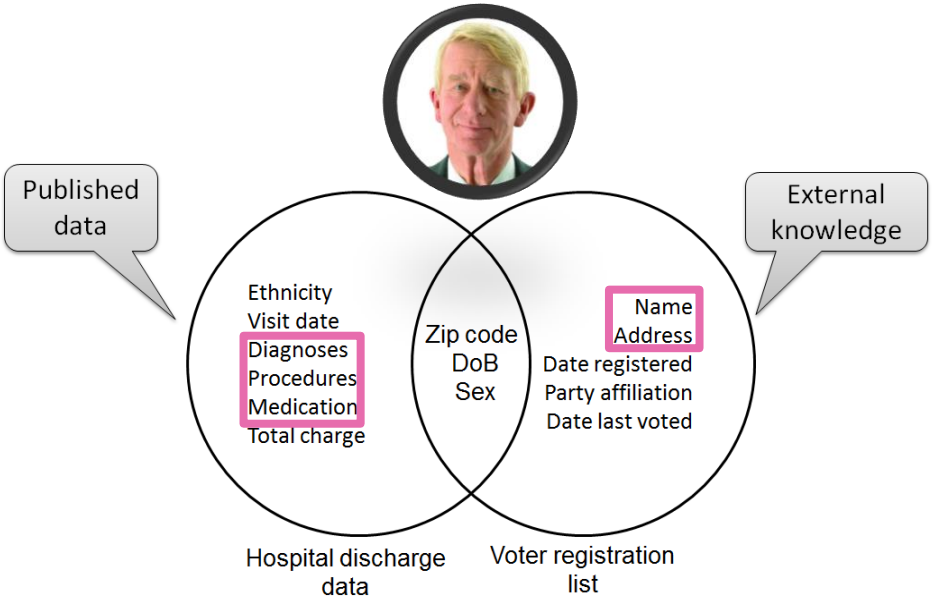
„[...] To determine whether a natural person is identifiable, **account should be taken of all the means reasonably likely to be used**, [..]to identify the natural person directly or indirectly [...]“

"[In doing so] all **objective factors, such as the costs of and the amount of time required** for identification, taking into consideration the **available technology at the time of the processing and technological developments** [...]"

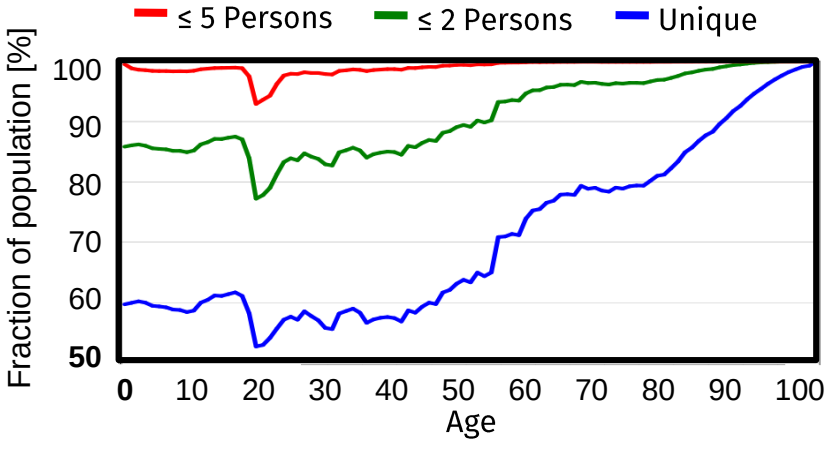
Source: Regulation (EU) 2016/679 of the European parliament and the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

Background: re-identification in 1997

Removing directly identifying attributes is not sufficient!



Around 87% of the U.S. population can be uniquely identified using a combination of 5-digit ZIP code, date of birth and sex



Source: Golle P. Revisiting the uniqueness of simple demographics in the US population. 5th ACM Workshop on Privacy in the Electronic Society, 2006, Sweeney L. Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000, Image by By Gary Johnson from Taos, NM - BillWeld5x7 (2), CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=49683363

Background: re-identification in 2019

Medical Data De-Identification Is Under Attack

Forbes

David Talby Forbes Councils Member
Forbes Technology Council COUNCIL POST | Paid Program
Innovation

POST WRITTEN BY

David Talby

PhD, MBA, CTO at Pacific AI. Making AI, big data and data science solve real-world problems in healthcare, life science and related fields.

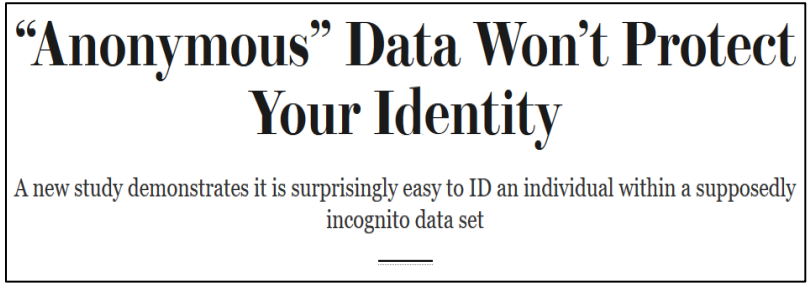
Forbes - Forbes Technology Council, 27.08.2019

The New York Times

Your Data Were 'Anonymized'? These Scientists Can Still Identify You

Computer scientists have developed an algorithm that can pick out almost any American in databases supposedly stripped of personal information.

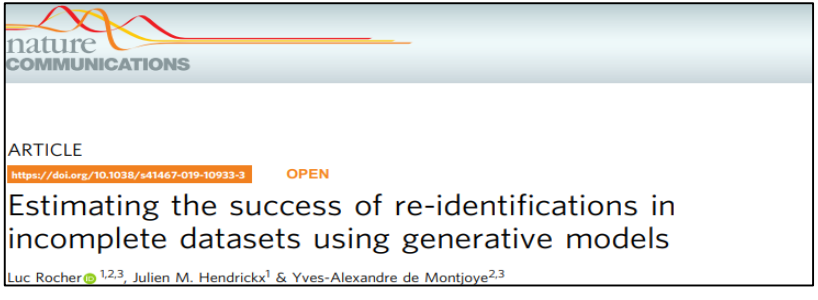
The New York Times, 23.07.2019



"Anonymous" Data Won't Protect Your Identity

A new study demonstrates it is surprisingly easy to ID an individual within a supposedly incognito data set

Scientific American, 23.07.2019



nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-10933-3 OPEN

Estimating the success of re-identifications in incomplete datasets using generative models

Luc Rocher[1,2,3], Julien M. Hendrickx[1] & Yves-Alexandre de Montjoye[2,3]

Nature Communications, 23.07.2019

"[...] we find that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes."

Background: Further examples of re-identification*

- **Demographic data** (Sweeney 1997; Golle 2006; El Emam 2008)
- **Diagnosis codes** (Loukides et al. 2010)
- **DNA (SNPs)** (Lin, Owen, & Altman 2004; Homer et al. 2008, Wang et al. 2009)
- **Pedigree structure** (Malin 2006)
- **Location visits** (Malin & Sweeney 2004, Golle & Partridge 2009)
- **Movie reviews** (Narayanan & Shmatikov 2008)
- **Search queries** (Barbaro & Zeller 2006)
- **Social network structure** (Backstrom et al. 2007, Narayanan & Shmatikov 2009)

*Based on: Bradley Malin. Challenges and Solutions for Data Privacy in Translational Research. 2011

Background: Technical perspective

- Processing of personal (input) data in such a way that anonymous (output) data is produced. Example:

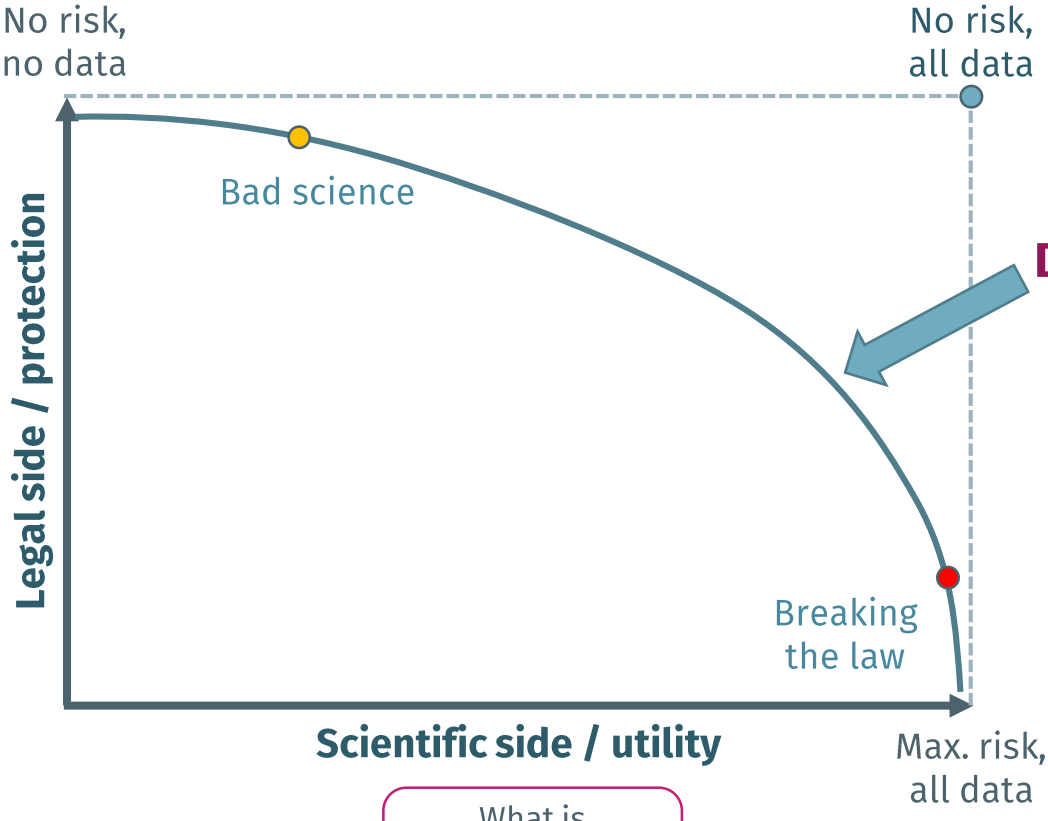
Alter	Geschlecht	PLZ	Gewicht	Diagnose
55	Männlich	81539	71	C25.0 Bösartige Neubildung des Pankreas - Pankreaskopf
76	Männlich	81675	80	C25.0 Bösartige Neubildung des Pankreas - Pankreaskopf
66	Männlich	81929	85	C25.0 Bösartige Neubildung des Pankreas - Pankreaskopf
81	Männlich	80802	79	C25.1 Bösartige Neubildung des Pankreas - Pankreaskörper
74	Männlich	81249	88	C25.2 Bösartige Neubildung des Pankreas - Pankreasschwanz
71	Weiblich	80335	69	C18.2 - Bösartige Neubildung des Kolons - Colon ascendens
64	Weiblich	80339	71	C18.4 - Bösartige Neubildung des Kolons - Colon transversum
69	Männlich	80637	75	C18.7 - Bösartige Neubildung des Kolons - Colon sigmoideum
55	Weiblich	80638	77	C18.7 - Bösartige Neubildung des Kolons - Colon sigmoideum
61	Männlich	81667	67	C18.7 - Bösartige Neubildung des Kolons - Colon sigmoideum

Alter	Geschlecht	PLZ	Gewicht	Diagnose
72,0	Männlich	81***	[80, 90[	C25.- Bösartige Neubildung des Pankreas
72,0	Männlich	81***	[80, 90[	C25.- Bösartige Neubildung des Pankreas
72,0	Männlich	81***	[80, 90[	C25.- Bösartige Neubildung des Pankreas
62,7	---	80***	[70, 80[	C18.- Bösartige Neubildung des Kolons
62,7	---	80***	[70, 80[	C18.- Bösartige Neubildung des Kolons
62,7	---	80***	[70, 80[	C18.- Bösartige Neubildung des Kolons

- Sampling
- Aggregation
- Suppression
- Masking
- Categorization
- Generalisation

Source: Fabian Prasser, Johanna Eicher, Helmut Spengler, Raffael Bild, Klaus A. Kuhn. Flexible Data Anonymization Using ARX — Current Status and Challenges Ahead. J Software Pract Exper 50, 7 (2020);1277-1304.

Background: Trade-offs



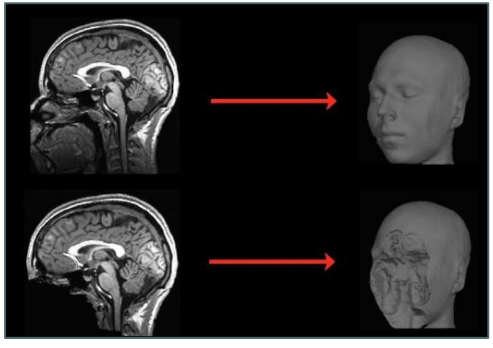
What is
- personal data?
- identification?
- acceptible risk?

What is
- planned use?
- requirements?

Source: Barth-Jones, Brussels Privacy Symposium, 2016

Background: A context-specific problem

- Purpose, recipient, types of data etc.



Source: https://surfer.nmr.mgh.harvard.edu/fswiki/mri_deface

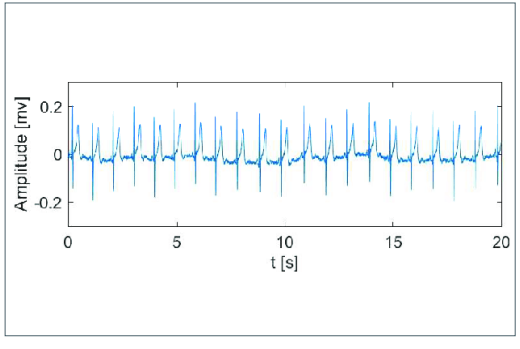
AUTOPSY REPORT - Final Anatomic Diagnosis

Dx: Sickle cell anemia with multiple red blood cell transfusions

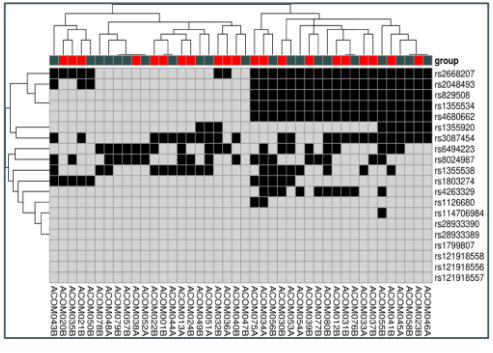
Cause of death per autopsy report (AU-01-23): Cirrhosis related to Hepatitis G

Mr. Herman Hesse is a 50 year old male, originally from Sri Lanka, who was diagnosed with sickle cell anemia at age 5. From the age of 7 to 11 1/2, he had several health complications and underwent a liver transplant at the Camelot Hospital Center mid-November 2016. He has been in good health and continued with normal daily activities until Dec. 2056, when he was brought to the Steppenwolf Clinic and admitted to the ICU. At that time, he was diagnosed with end-stage renal disease. He responded well to hemodialysis for about a year per his wife, Hermine Mozart. A few months later he began to experience chronic pain in his left hip and was referred to Dr. Goethe at the Everyone's Well Pain Management Center. On October 1st, 2057, he was re-admitted to the Steppenwolf Clinic and quickly transferred to the ICU. Due to his declining health, the patient's wife met with an ethics consultant and decided to withdraw medical services and provide comfort measures only. The patient expired on October 6th, 2057. A limited autopsy was performed on the sixth of October at 3:00pm.

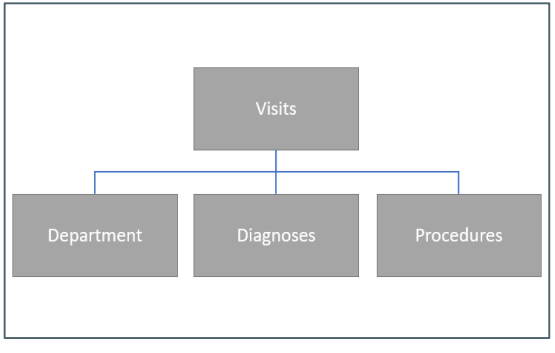
Source : https://scrubber.nlm.nih.gov/



Source : https://doi.org/10.1109/MeMeA.2018.8438751



Source : https://doi.org/10.2147/CCID.S176842



Source : https://www.g-drg.de/Datenlieferung_gem._21_KHEntgG

Onset of exposure	Yes	No	Total
20+ years***	339	53	392
0–19 years***	203	522	725
Total	542	575	1,117

Source : https://doi.org/10.1080/10937404.2012.678766

Background: Tools for structured data

- Automatic or semi-automatic procedures for solving the risk/utility optimization problem.
- Can support various mathematical and statistical models for quantifying risks and data utility (i.e. independent of a specific law or interpretation).
- Mature open source tools

  sdcMicro, sdcGUI and sdcTable
  - Packages for R statistics environment for individual-level data and statistical tables. Semi-automated process. Selected functions.

  ARX
  - Java programming library and stand-alone tool for individual level data. More automated process. Comprehensive set of features.

2. Threats and protection methods

Recap: Technical perspective

- Processing of personal (input) data in such a way that anonymous (output) data is produced. Example:

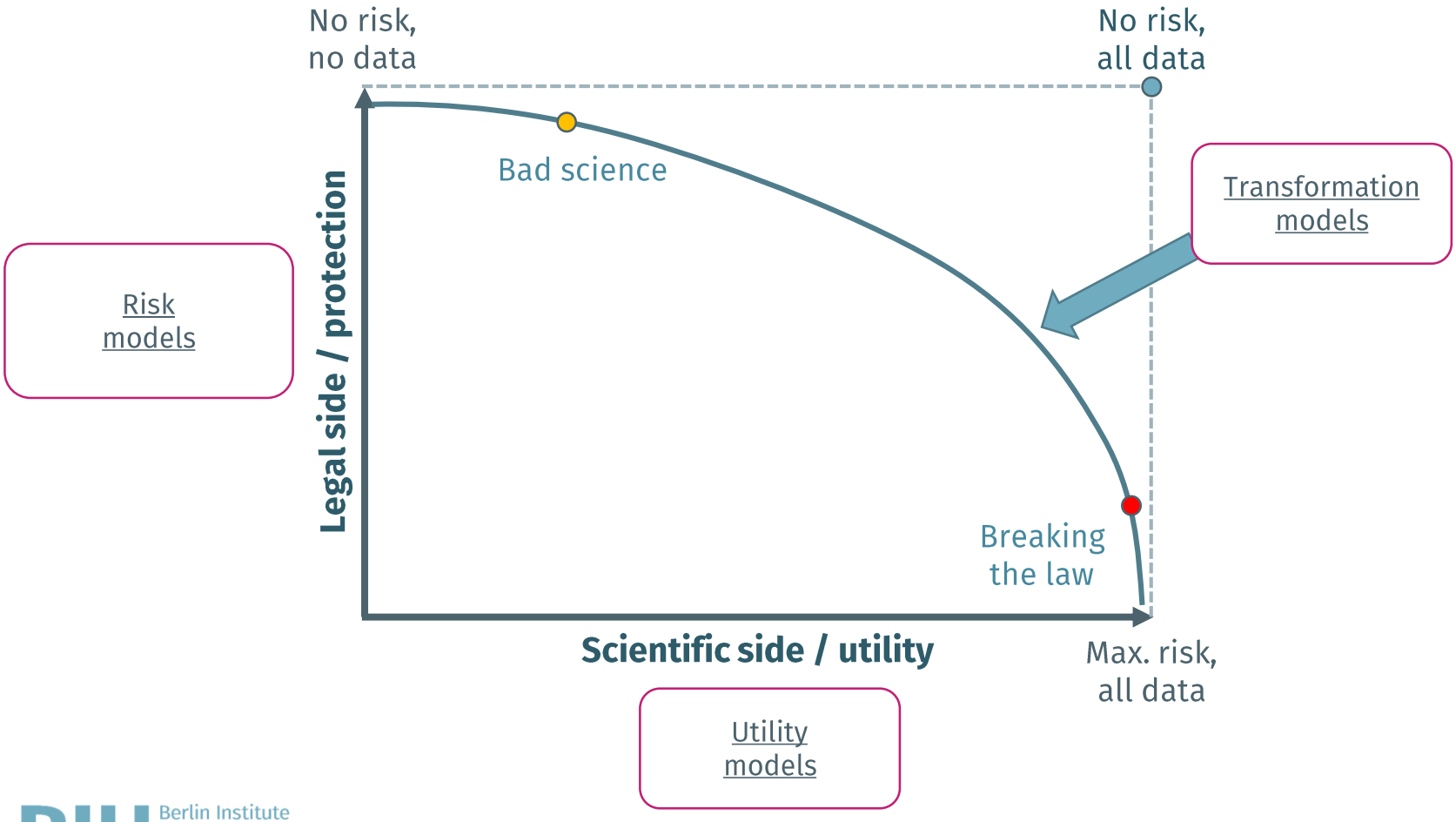
Alter	Geschlecht	PLZ	Gewicht	Diagnose
55	Männlich	81539	71	C25.0 Bösartige Neubildung des Pankreas - Pankreaskopf
76	Männlich	81675	80	C25.0 Bösartige Neubildung des Pankreas - Pankreaskopf
66	Männlich	81929	85	C25.0 Bösartige Neubildung des Pankreas - Pankreaskopf
81	Männlich	80802	79	C25.1 Bösartige Neubildung des Pankreas - Pankreaskörper
74	Männlich	81249	88	C25.2 Bösartige Neubildung des Pankreas - Pankreasschwanz
71	Weiblich	80335	69	C18.2 - Bösartige Neubildung des Kolons - Colon ascendens
64	Weiblich	80339	71	C18.4 - Bösartige Neubildung des Kolons - Colon transversum
69	Männlich	80637	75	C18.7 - Bösartige Neubildung des Kolons - Colon sigmoideum
55	Weiblich	80638	77	C18.7 - Bösartige Neubildung des Kolons - Colon sigmoideum
61	Männlich	81667	67	C18.7 - Bösartige Neubildung des Kolons - Colon sigmoideum

Alter	Geschlecht	PLZ	Gewicht	Diagnose
72,0	Männlich	81***	[80, 90[	C25.- Bösartige Neubildung des Pankreas
72,0	Männlich	81***	[80, 90[	C25.- Bösartige Neubildung des Pankreas
72,0	Männlich	81***	[80, 90[	C25.- Bösartige Neubildung des Pankreas
62,7	---	80***	[70, 80[	C18.- Bösartige Neubildung des Kolons
62,7	---	80***	[70, 80[	C18.- Bösartige Neubildung des Kolons
62,7	---	80***	[70, 80[	C18.- Bösartige Neubildung des Kolons

- Sampling
- Aggregation
- Suppression
- Masking
- Categorization
- Generalisation

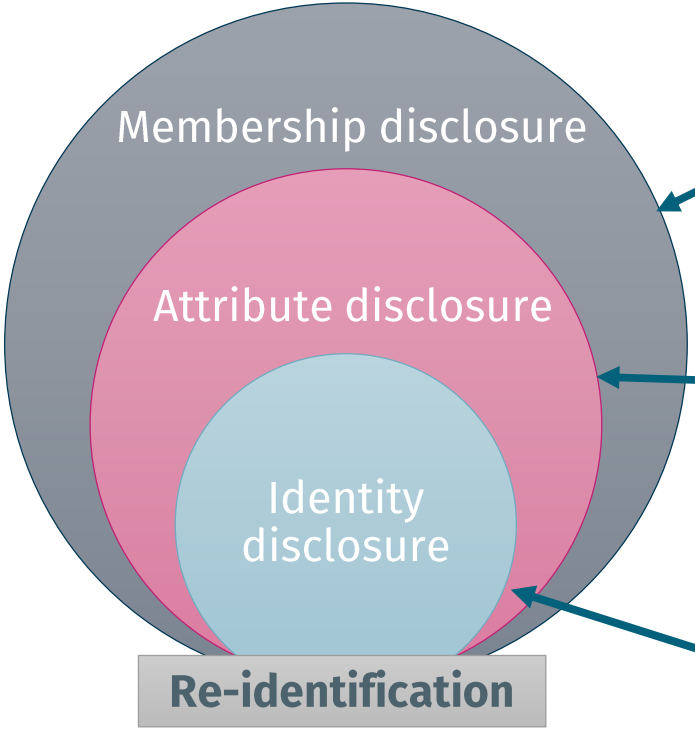
Source: Fabian Prasser, Johanna Eicher, Helmut Spengler, Raffael Bild, Klaus A. Kuhn. Flexible Data Anonymization Using ARX — Current Status and Challenges Ahead. J Software Pract Exper 50, 7 (2020);1277-1304.

Complexity: axes



Source: Barth-Jones, Brussels Privacy Symposium, 2016

Complexity: types of attacks

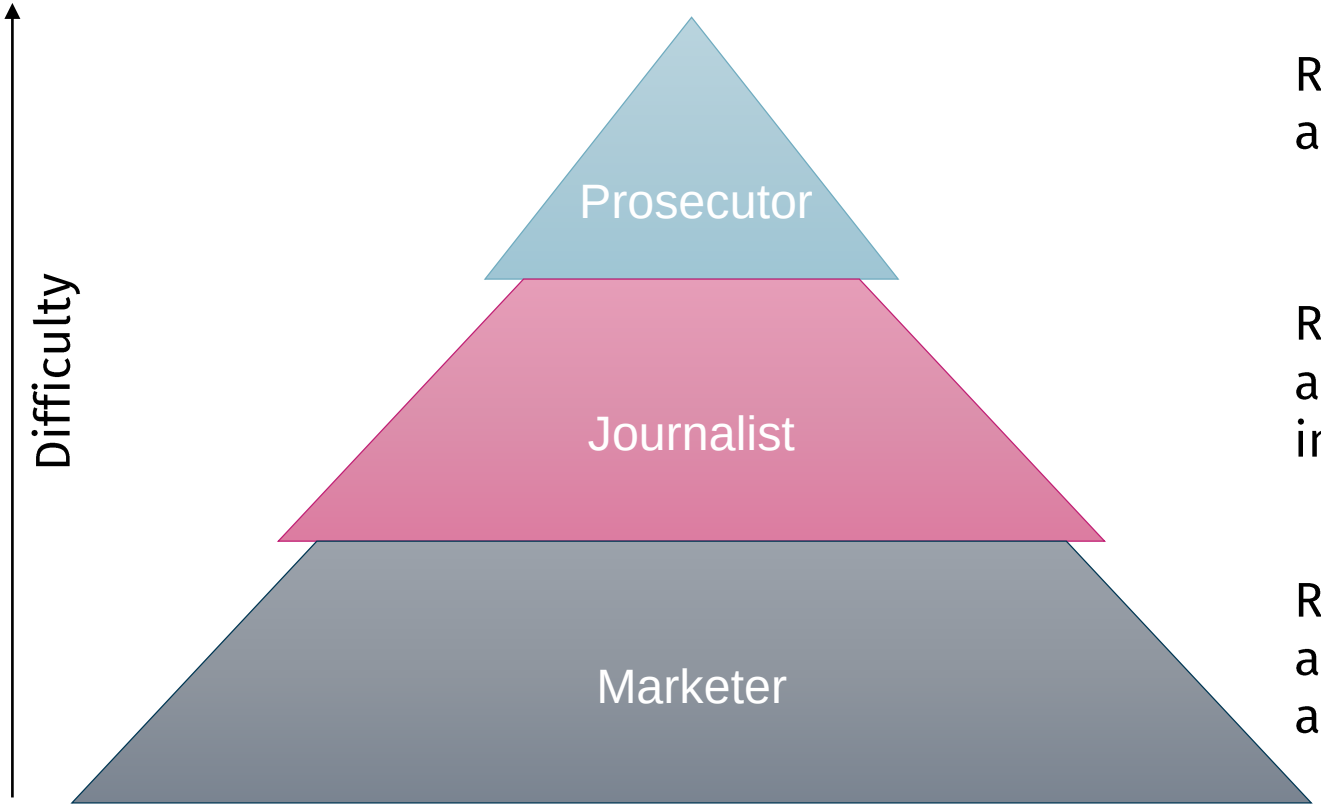


Is data about an individual contained?

Can additional information be inferred about individuals?

Can an individual be associated to a specific record?

Complexity: types of re-identification

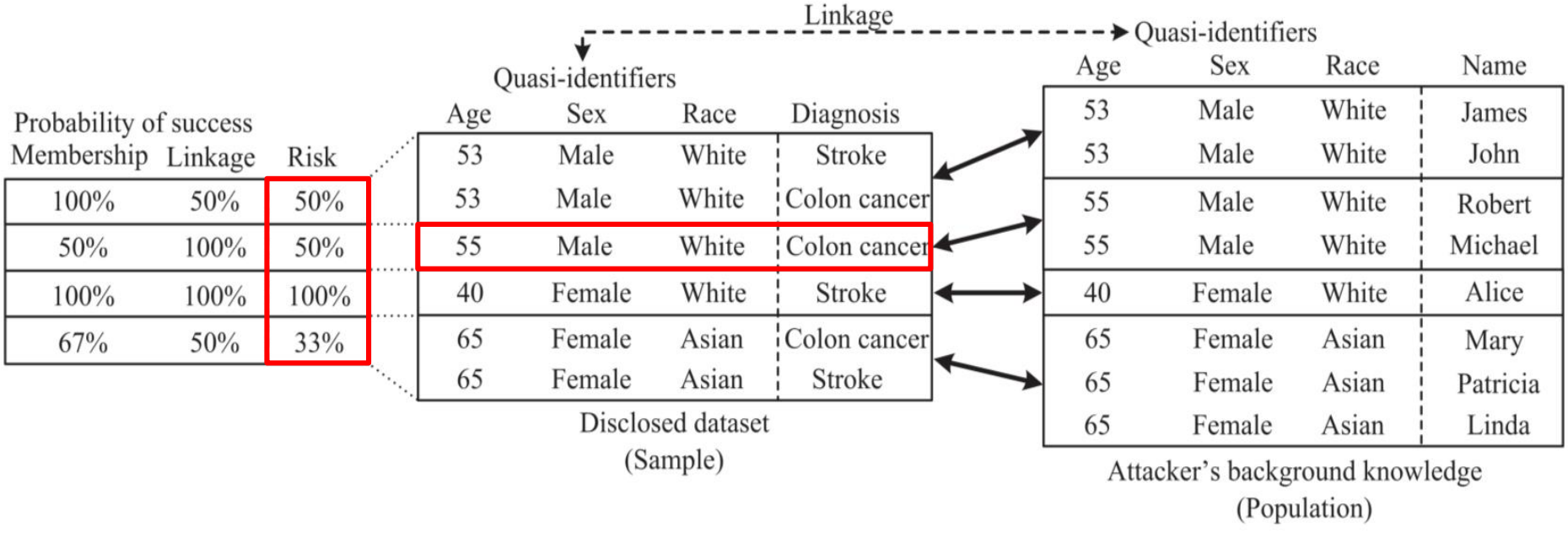


Re-identification of a specific individual

Re-identification of an unspecific individual

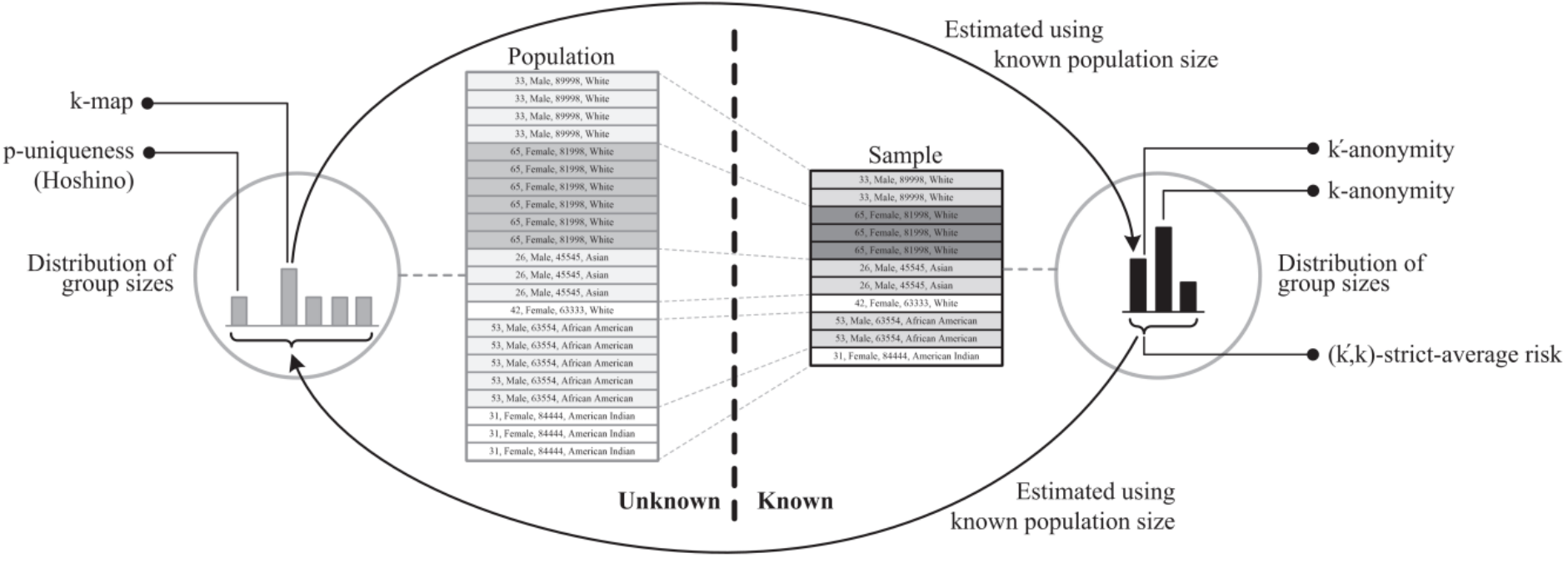
Re-identification of as many individuals as possible

Example: calculation of re-identification risks



Source: Fabian Prasser, Florian Kohlmayer, Klaus A. Kuhn. The Importance of Context: Risk-Based De-Identification of Biomedical Data. Methods of Information in Medicine, Schattauer, June 2016.

Example: estimation of re-identification risks



Source: Fabian Prasser, Florian Kohlmayer, Klaus A. Kuhn. The Importance of Context: Risk-Based De-Identification of Biomedical Data. Methods of Information in Medicine, Schattauer, June 2016.

A new perspective: differential privacy

Randomized "mechanism"

Data

Plausible deniability →

Basically the same

ε

Randomized "mechanism"

Data

interactive or non-interactive

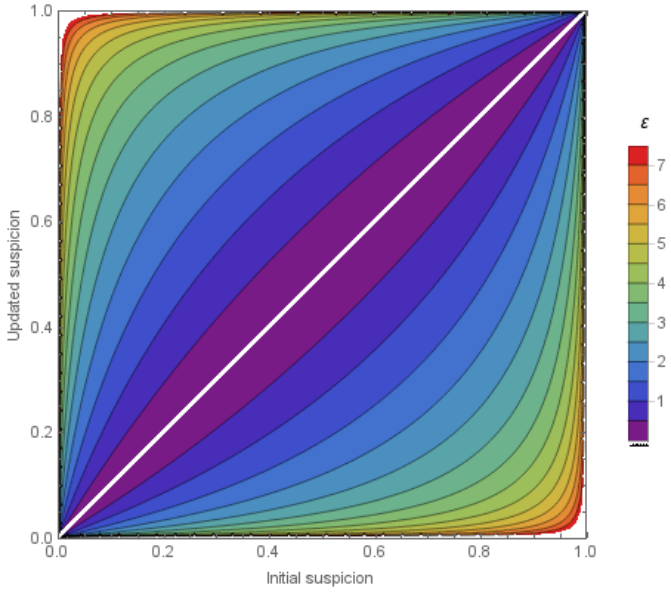
Source: https://desfontain.es/privacy/differential-privacy-in-more-detail.html

Differential privacy: pros and cons

- No need to make assumptions about attacks
  - Protects any kind of information about any individual
  - Works regardless of the attacker's background knowledge
- Risk can be quantified, e.g. “membership”
- Composition of mechanisms

But

- Many mechanisms are not truthful
- Differential Privacy is not very intuitive and often difficult to communicate



Source: https://desfontain.es/privacy/differential-privacy-in-more-detail.html

3. Anonymization of analysis results

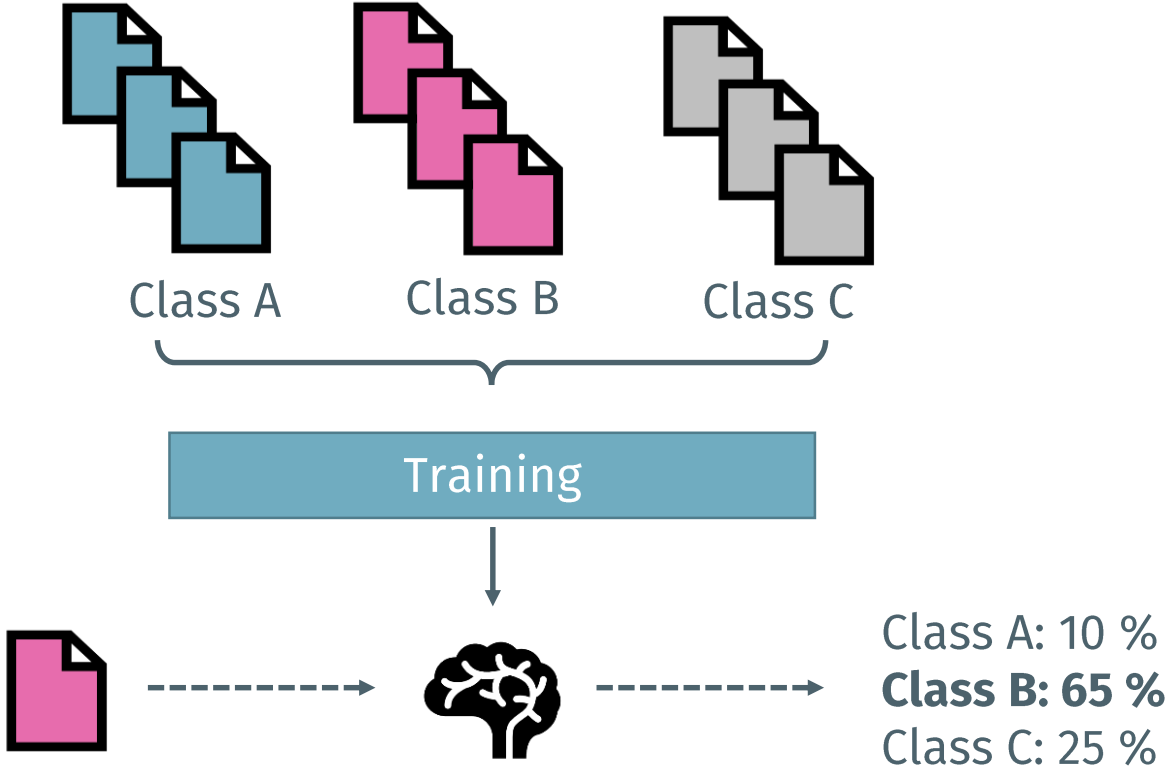
GDPR: Recital 162

Also the output of data analyses must be protected!

- „Where personal data are processed for statistical purposes, this Regulation should apply to that processing. […]“
- „[…] Statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results. […]“
- „[…] Those statistical results may further be used for different purposes, **including a scientific research purpose**. […]“
- „[…] The statistical purpose **implies that the result of processing for statistical purposes is not personal data, but aggregate data**, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person. […]“

Source: Regulation (EU) 2016/679 of the European parliament and the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

Example: What can be learned from classification models?



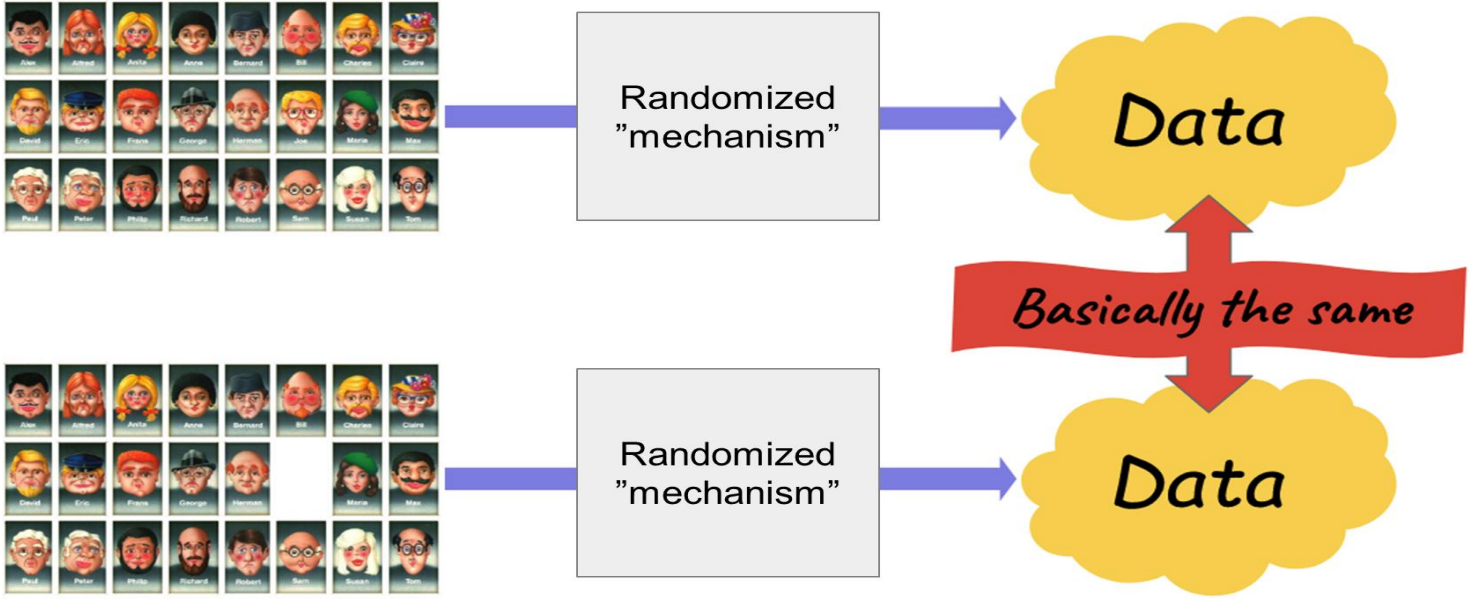
Quelle: Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017, May). Membership inference attacks against machine learning models. In 2017 IEEE Symposium on Security and Privacy (SP) (pp. 3-18). IEEE.

Example: Attack vectors on classification models

- **Membership disclosure**
  - For inputs that can be classified with a high accuracy it is more likely that they have been used to train the model
  - Shadow model attacks
- **Attribute disclosure**
  - The output of a model can be used to draw conclusions about input data if some features and the expected prediction are known
- **Data leakage**
  - For example in text mining, where tokens might be encoded into models

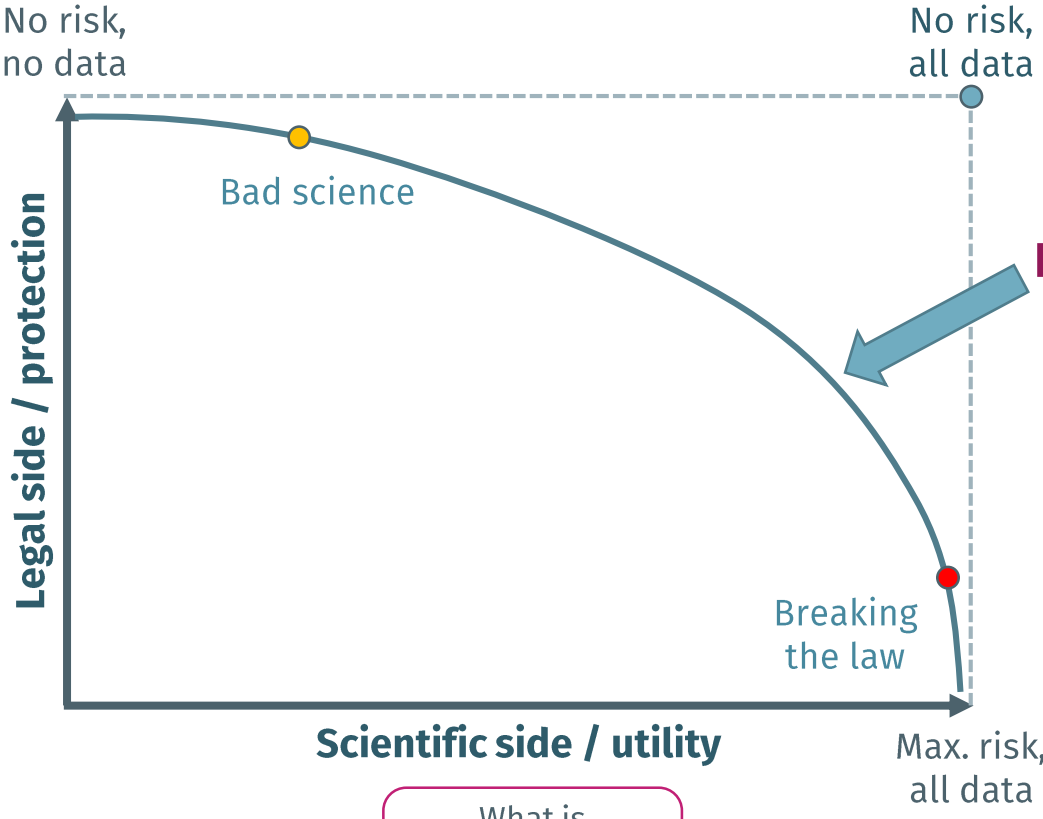
Protection: TensorFlow Privacy

- Implements the privacy model Differential Privacy into TensorFlow



Quellen: McMahan HB, Andrew G. A General Approach to Adding Differential Privacy to Iterative Training Procedures. arXiv preprint arXiv:1812.06210. 2018
https://medium.com/tensorflow/introducing-tensorflow-privacy-learning-with-differential-privacy-for-training-data-b143c5e801b6 Accessed: 29.3.2019
Grafik: https://desfontain.es/privacy/differential-privacy-in-more-detail.html Accessed: 28.03.2019

Recap: Trade-offs

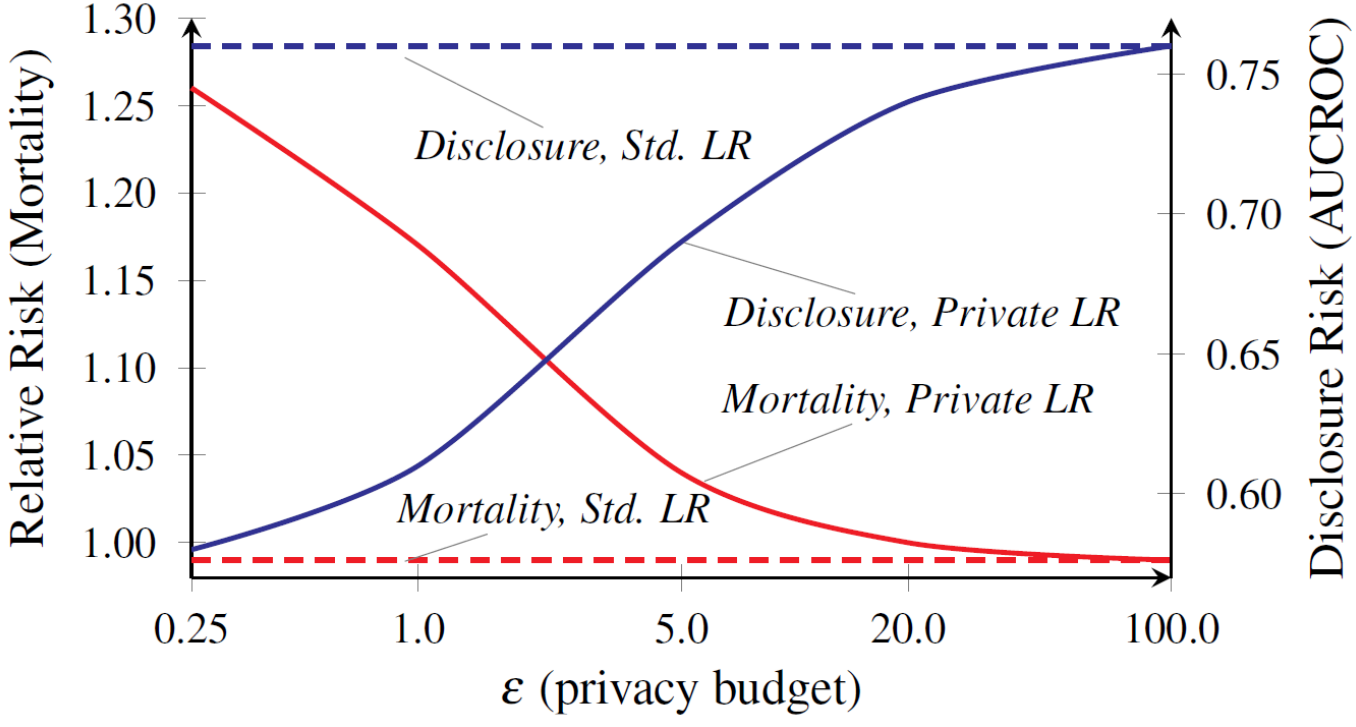


What is
- personal data?
- identification?
- acceptible risk?

What is
- planned use?
- requirements?

Source: Barth-Jones, Brussels Privacy Symposium, 2016

Example: Dosage of Warfarin

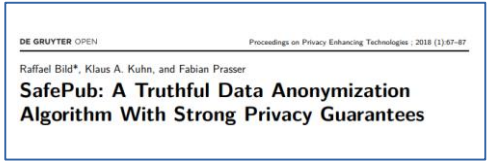


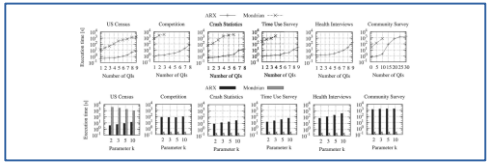
Source: Fredrikson, M., Lantz, E., Jha, S., Lin, S., Page, D., & Ristenpart, T. (2014). Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing. In 23rd USENIX Security Symposium (pp. 17-32).

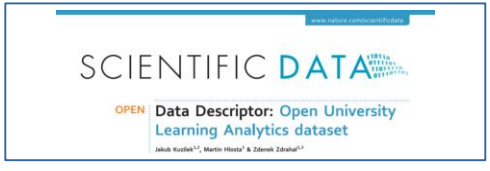
4. ARX Data Anonymization Tool

ARX: Features and applications

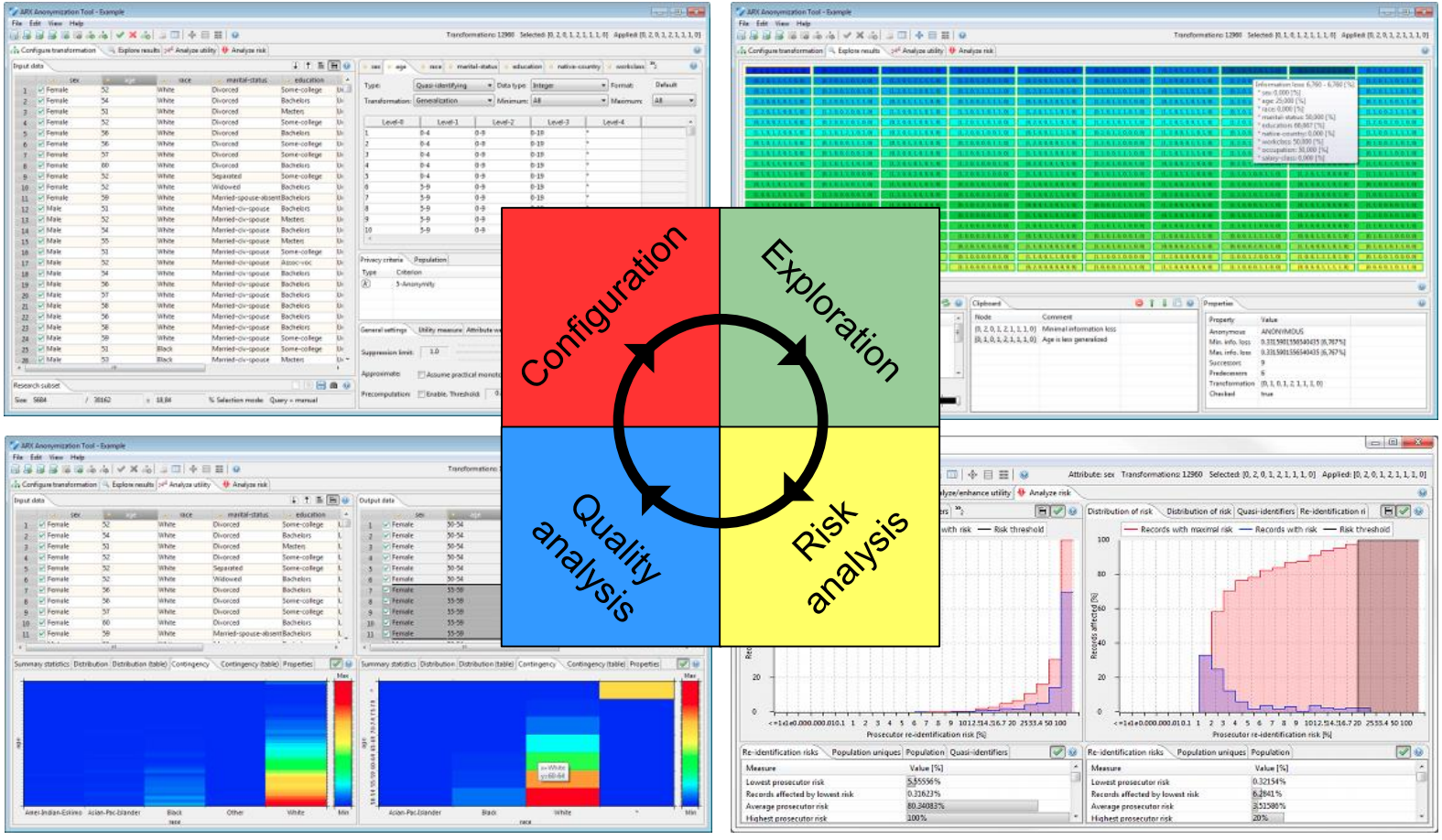
- **Comprehensive feature set**: „traditional“ approaches, Differential Privacy, game-theoretic methods, privacy-preserving machine learning.
- **Quite scalable**: Significantly outperforms related tools, used to anonymise datasets with billions of records.
- **Graphical tool**: Used in education and training by commercial and public institutions in several countries.
- **Wide range of applications**: Creation of open datasets and used to build anonymisation pipelines in several domains, e.g. by telecom providers, health insurances.
- **Industry friendly**: Integrated into several commercial products, core algorithms adopted by SAP HANA.
- **Open source:** More than 50.000 downloads.





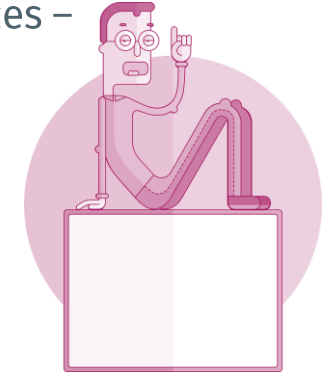


ARX: Graphical frontend

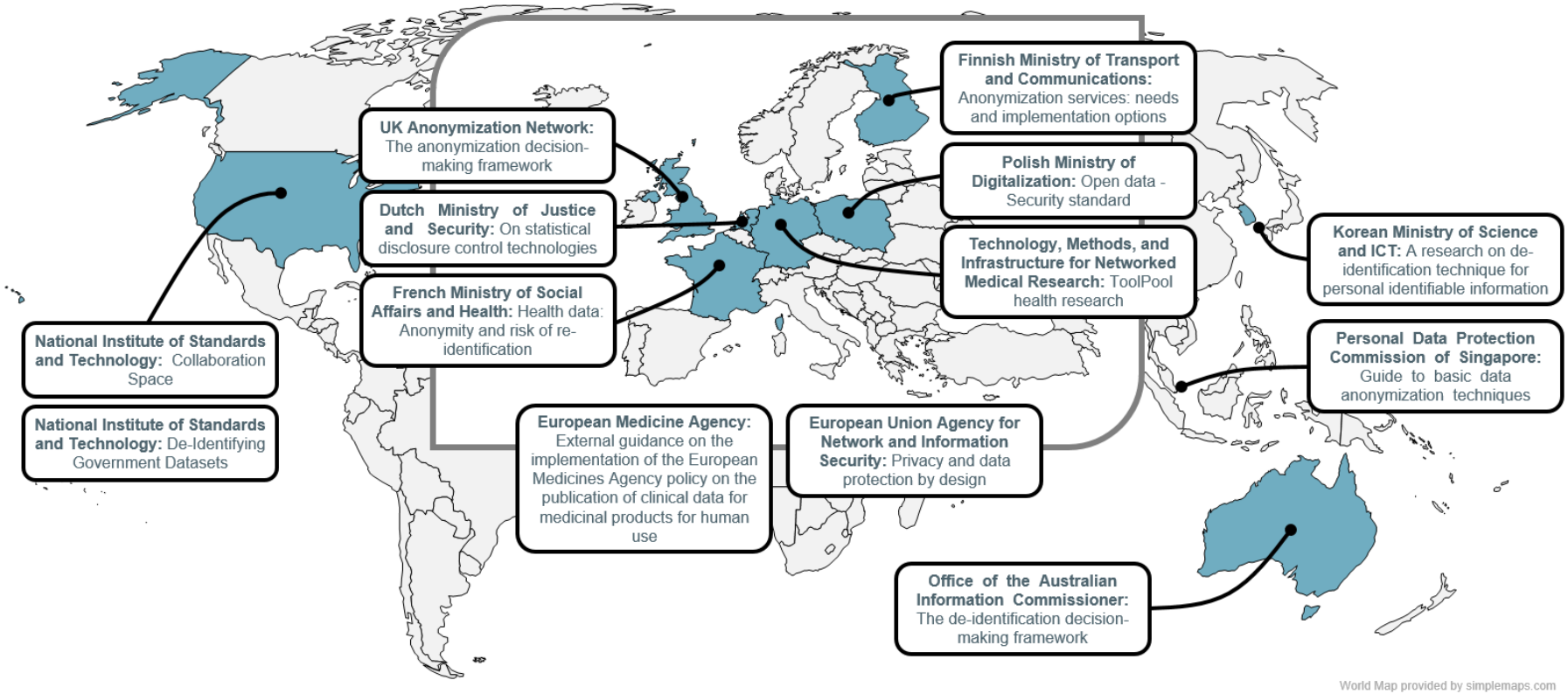


Examples of guidelines mentioning ARX (1)

- European Medicines Agency. EMA/90915/2016 – external guidance on the implementation of the European medicines agency policy on the publication of clinical data for medicinal products for human use; 2018.
- European Union Agency for Network and Information Security. Privacy and data protection by design; 2015.
- UKAN. The anonymisation decision-making framework; 2016.
- Office of the Australian Information Commissioner. The de-identification decision-making framework; 2017.
- French Ministry of Solidarity and Health. Health data: anonymity and risk of re-identification; 2015.
- Finnish Ministry of Transport and Communications. Anonymization services – requirements and implementation options; 2017.
- Personal Data Protection Commission of Singapore. Guide to basic data anonymisation techniques; 2018.
- Polish Ministry of Digitalization. Open data - Security standard; 2018.
- Dutch Ministry of Justice and Security. On statistical disclosure control technologies; 2018.
- Korean Ministry of Science and ICT. A research on de-identification technique for personal identifiable information; 2016.



Examples of guidelines mentioning ARX (2)



5. Real-World Examples

Example: Anonymisation pipelines for the LEOSS registry

- LEOSS: A European registry capturing the clinical course of SARS-CoV-2 infected patients (https://leoss.net) established at University of Cologne
  - No informed consent necessary (anonymous reports).
  - Retrospective documentation after discharge / death.
  - All hospitalized patients including children eligible.
  - Immediate start after verification.
- Open Science approach
  - Registry hosted in a secure environment in Cologne.
  - Anonymous data is shared with researchers and the public.
  - Additional anonymisation procedures have been implemented for this purpose.

LEOSS

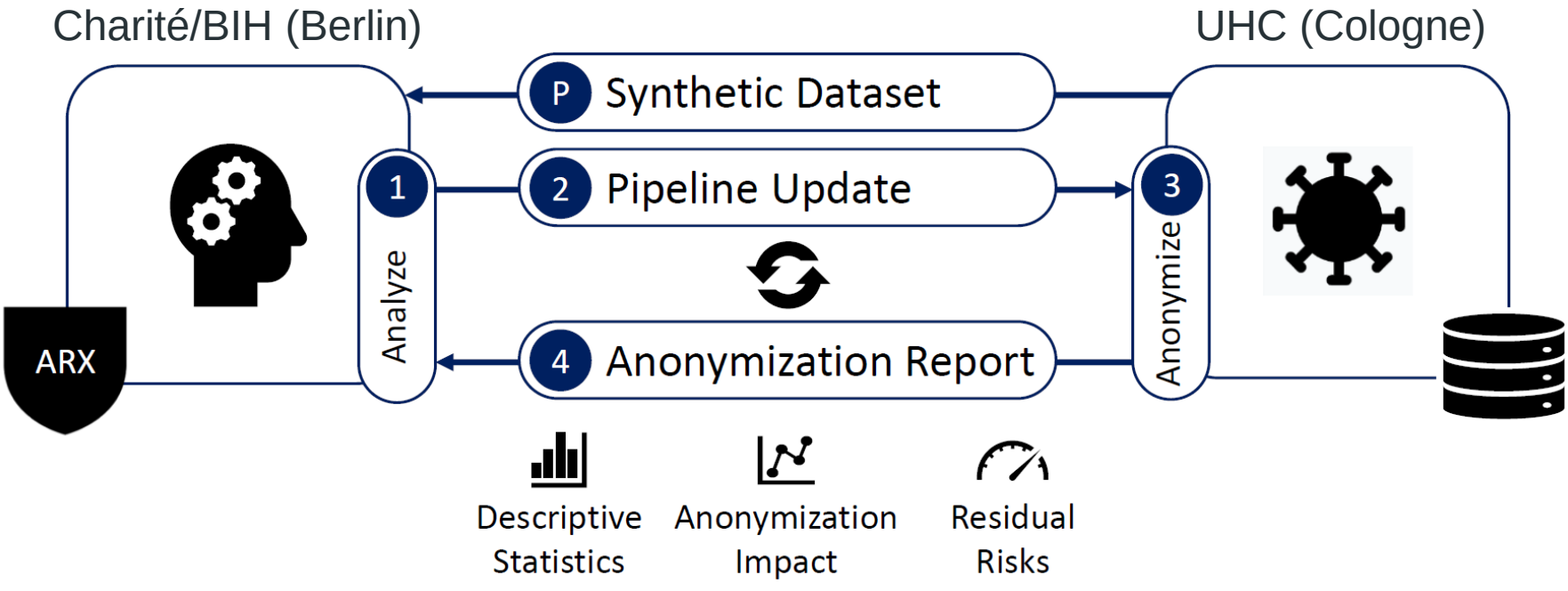
Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

LEOSS: Overview

- Two types of datasets
  - Public Use File with 16 variables available without restrictions.
  - Scientific Use Files with ≤605 variables available under data use contracts.
- Two types of pipelines, built with ARX
  - Two stages for the Public Use File
  - Ten stages for the Scientific Use File
- Both pipelines were developed without access to primary data in close cooperation with the LEOSS Core Team in Cologne.

Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

LEOSS: Development process



→ Seven iterations over several weeks

Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

LEOSS: Approach for the Public Use File (1)

(1) Qualitative risk assessment

- Compared data to “risky” variables mentioned in laws and guidelines.
  - Low risk already according to this initial assessment.
- Additionally, assessed the risk of identification associated with individual variables following a methodology proposed by Malin et al.*
  - Replicability, availability, distinguishability categorized into low, medium or high.
  - Variables above threshold considered potentially identifying.

(2) Quantitative risk assessment

- Followed recommendations from the Opinion on Anonymisation Methods by the Article 29 Data Protection Working Party (today: European Data Protection Board):
  - Singling out: the possibility to isolate some or all records which identify an individual in the dataset.
  - Linkability: the ability to link, at least, two records concerning the same data subject or a group of data subjects.
  - Inference: the possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes.

* Malin, B., Loukides, G., Benitez, K. & Clayton, E. W. Identifiability in biobanks: models, measures, and mitigation strategies. Hum. Genet. 130, 383–392 (2011).

LEOSS: Approach for the Public Use File (2)

(3) Formal anonymization process

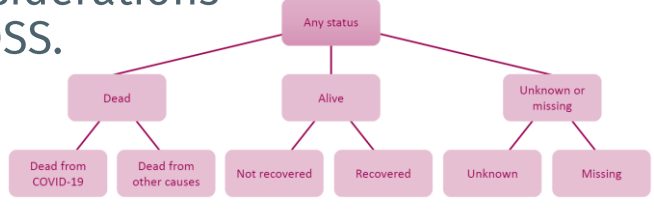
- Generalization and record suppression to mitigate risks highlighted by the Opinion.
- Prevented singling out and linkability by reducing the uniqueness of all possible combinations of potentially identifying variables (k-anonymity).
- Prevented inference by ensuring that the distribution of medical data within groups of indistinguishable records is not too different from the distribution in the overall dataset (t-closeness).
- Static generalization scheme and withholding of records to ensure that protection holds also when data is updated repeatedly.

(4) Extensive documentation

- Entire development process and underlying considerations are documented in detail. Pipeline released as OSS.

(5) Continuous monitoring

- Repeated evaluation of data utility.



Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

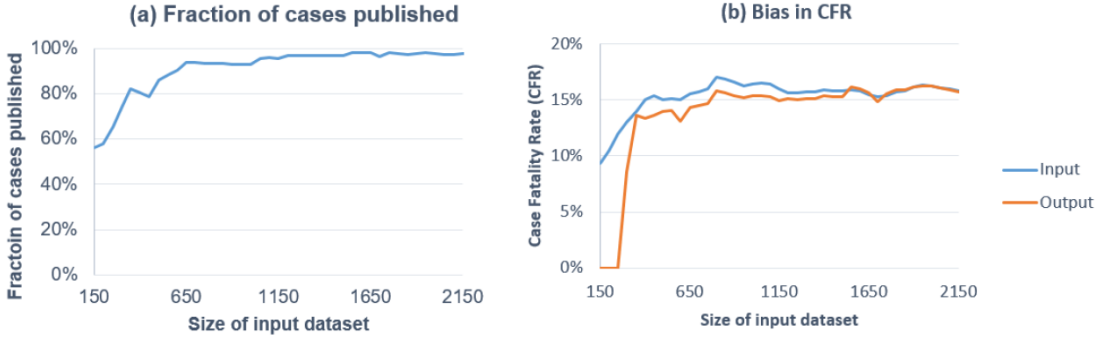
LEOSS: Result

Variable	Description
Age at diagnosis	Age of patient at time of diagnosis
Gender	Sex of patient
Month first diagnosis	Month of first confirmed diagnosis of COVID-19
Year first diagnosis	Year of first confirmed diagnosis of COVID-19
Uncomplicated phase	Indicates whether the patient has been through the uncomplicated phase of COVID-19
Complicated phase	Indicates whether the patient has been through the complicated phase of COVID-19
Critical phase	Indicates whether the patient has been through the critical phase of COVID-19
Recovery phase	Indicates whether the patient has been through the recovery phase of COVID-19
Vasopressors in complicated phase	Indicates whether vasopressors where used in the complicated phase
Vasopressors in critical phase	Indicates whether vasopressors where used in the critical phase
Invasive ventilation in critical phase	Indicates whether invasive ventilation was used in the critical phase
Superinfection in uncomplicated phase	Type of (if any) superinfection in uncomplicated phase
Superinfection in complicated phase	Type of (if any) superinfection in complicated phase
Superinfection in critical phase	Type of (if any) superinfection in critical phase
Symptoms in recovery phase	Symptoms (if any) in recovery phase
Last known patient status	Last known status

Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

LEOSS: Evaluation (1)

- Pipeline based on the principle of “hiding in the crowd”
  - Anonymity is achieved by making sure that each record does not differ significantly from a larger group of records.
  - Counter-intuitive property: the greater the number of individuals included in the registry, the less information has to be removed to achieve the required degree of protection.
- Example: records released and case fatality rate

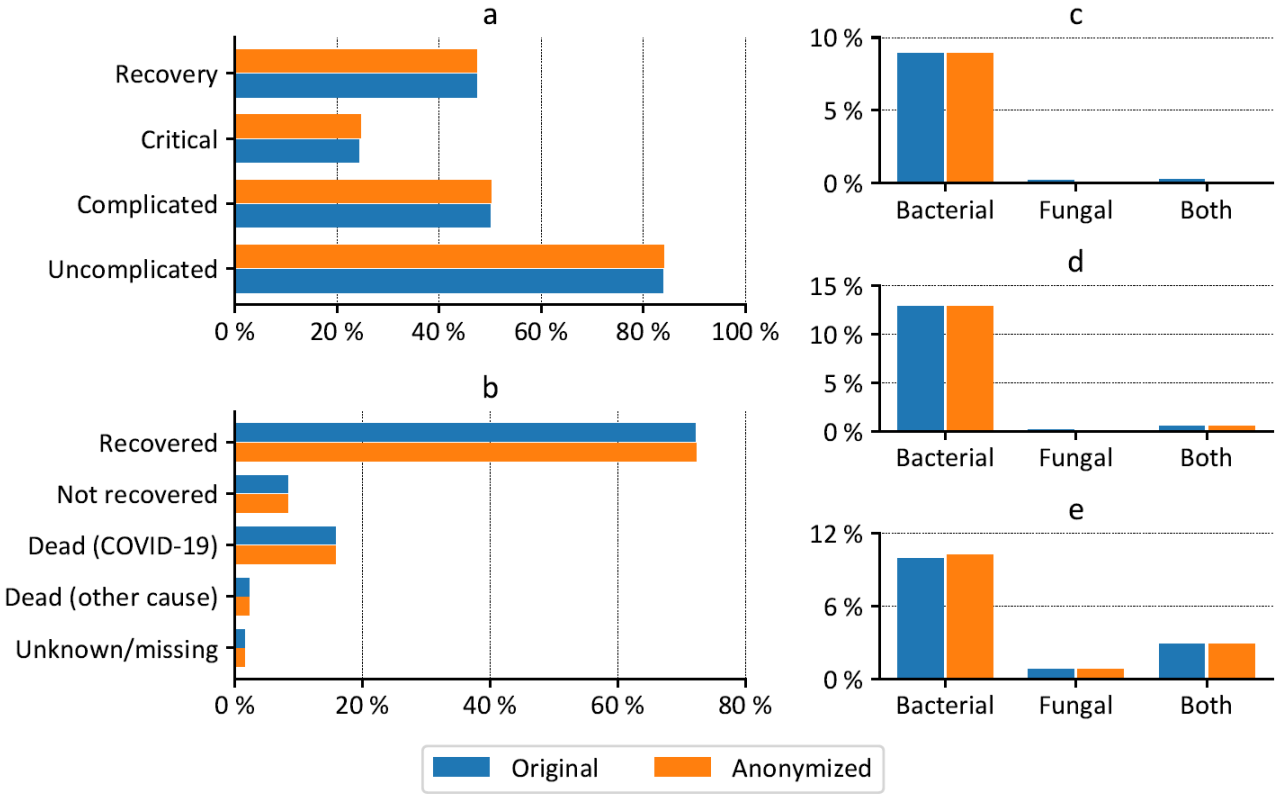


→ Negligible impact on data utility!

Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

LEOSS: Evaluation (2)

- Example: descriptive statistics



Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

LEOSS: Summary

- Eight additional pipeline stages implement transformations for various modules of the Scientific Use File. Examples:
  - Categorizing metric variables.
  - Making timestamps relative.
  - Grouping or suppressing sensitive variables.

  → Modules and stages can be activated dynamically to adjust to needs of different scientific / medical domains.
- Overall approach
  - Context-specific: adopted to the concrete dataset.
  - Multiple layers of safeguards: qualitative + quantitative methods.
  - Reliance on recommendations from laws and guidelines.
  - Risk-based approach requires thorough documentation.

Jakob CE, Kohlmayer F, Meurers T, Vehreschild JJ, Prasser F. Design and evaluation of a data anonymization pipeline to promote Open Science on COVID-19. Scientific data. 2020 Dec 10;7(1):1-0.

Thank you for your attention!

Prof. Dr. Fabian Prasser

Medical Informatics Group
Berlin Institute of Health @
Charité – Universitätsmedizin Berlin

https://www.bihealth.org/de/forschung/arbeitsgruppen/fabian-prasser/

Berlin Institute of Health
Charité & MDC

Aus Forschung wird Gesundheit