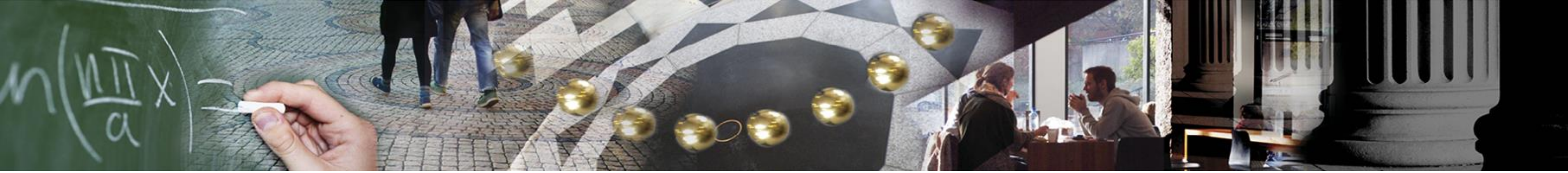


UiO : Department of Informatics
University of Oslo

Nils Gruschka

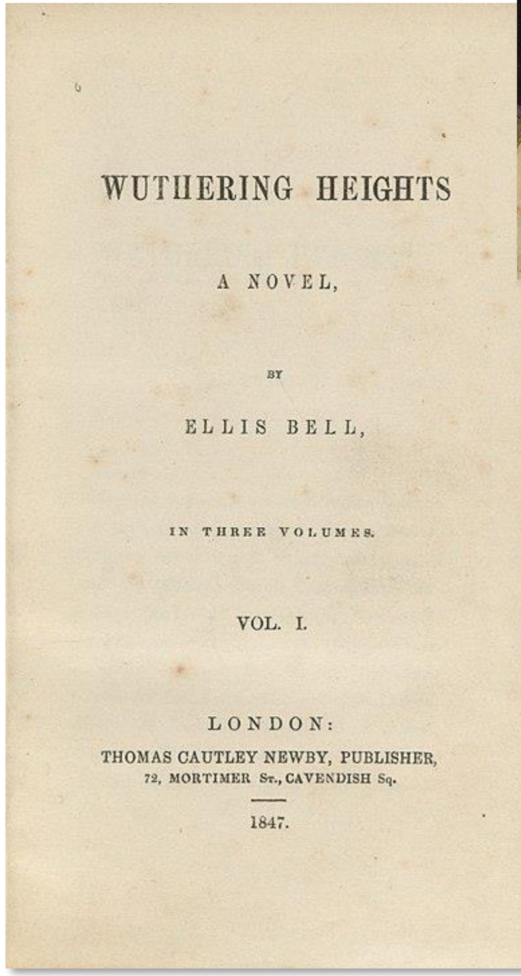
Pseudonymisation

Definitions

- *A pseudonym is an identifier of a subject other than one of the subject's real names.*

(Pfitzmann & Hansen, 2010)

Examples of Pseudonyms



WUTHERING HEIGHTS

A NOVEL,

BY

ELLIS BELL,

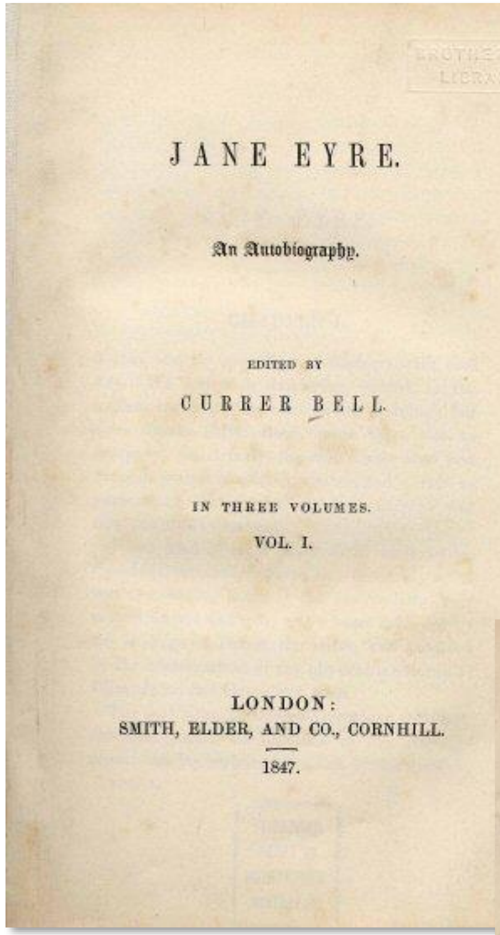
IN THREE VOLUMES.

VOL. I.

LONDON:
THOMAS CAUTLEY NEWBY, PUBLISHER,
72, MORTIMER St., CAVENDISH Sq.

1847.

Emily Jane Brontë



JANE EYRE.

An Autobiography.

EDITED BY

CURRER BELL.

IN THREE VOLUMES.

VOL. I.

LONDON:
SMITH, ELDER, AND CO., CORNHILL.

1847.

Charlotte Brontë

Joanne Rowling

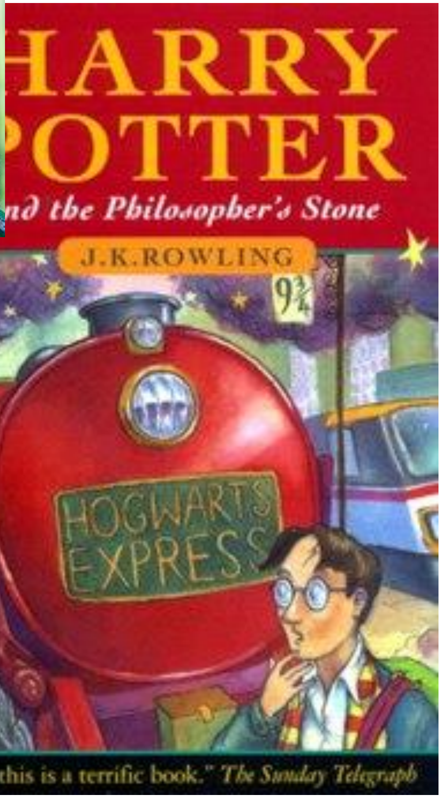

HARRY POTTER
and the Philosopher's Stone
J.K.ROWLING
9¾
HOGWARTS EXPRESS
"...this is a terrific book." The Sunday Telegraph

Image source: Wikipedia

Definitions

- *'Pseudonymisation' means the processing of personal data in such a manner that the personal data* ***can no longer be attributed to a specific data subject without the use of additional information****, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.*

(Regulation (EU) 2016/679, Art. 4(5))

Examples of Pseudonyms

- Medical / pharmaceutical research

ID 572345572

Patient	Disease	Age	Gender	Physic
…	…	…	…	
572345572	Influenza	57	male	Aspirin
…	…	…	…	

Image source: Wikipedia

Pseudonymisation Architectures (1)

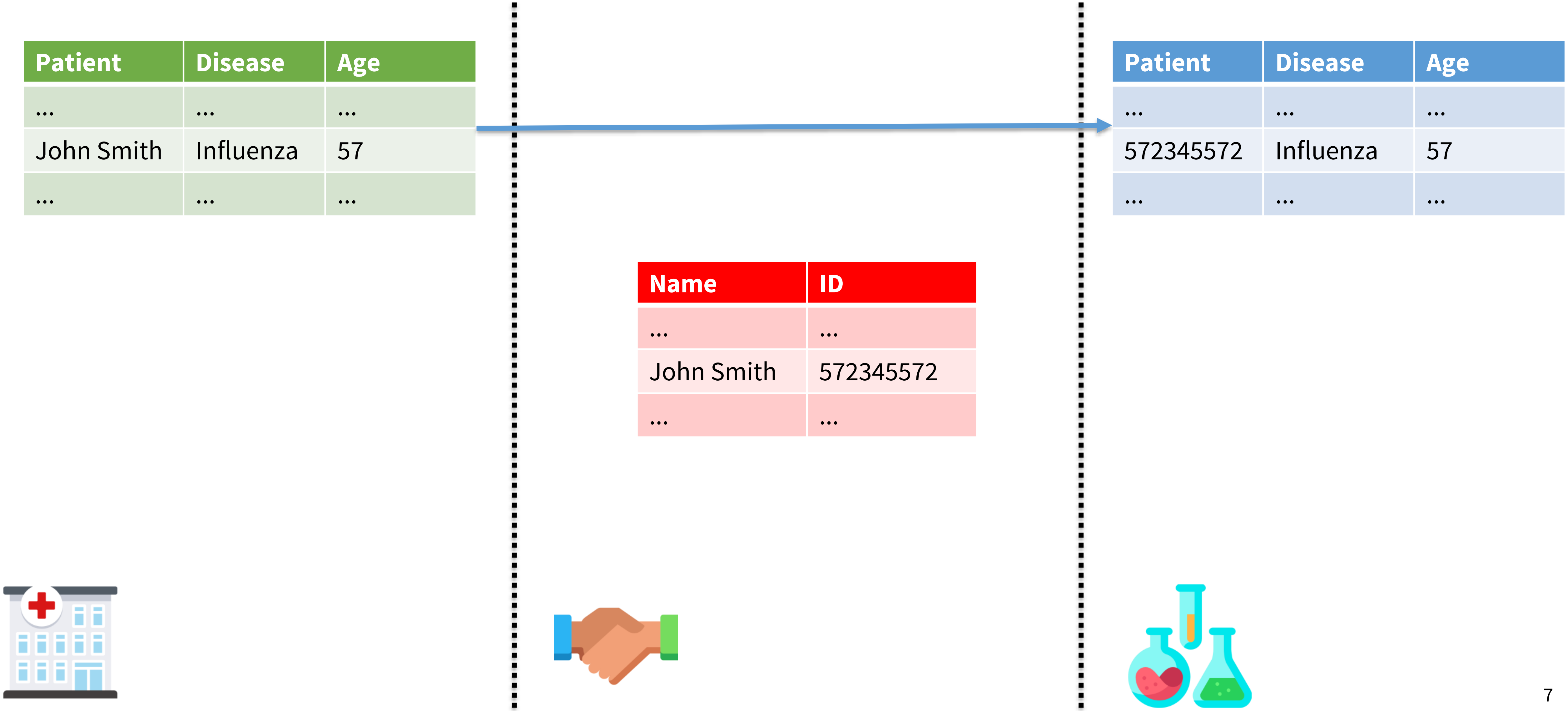
Patient	Disease	Age
…	…	…
John Smith	Influenza	57
…	…	…

Name	ID
…	…
John Smith	572345572
…	…

Patient	Disease	Age
…	…	…
572345572	Influenza	57
…	…	…

Icons made by Freepik from www.flaticon.com

Pseudonymisation Architectures (2)



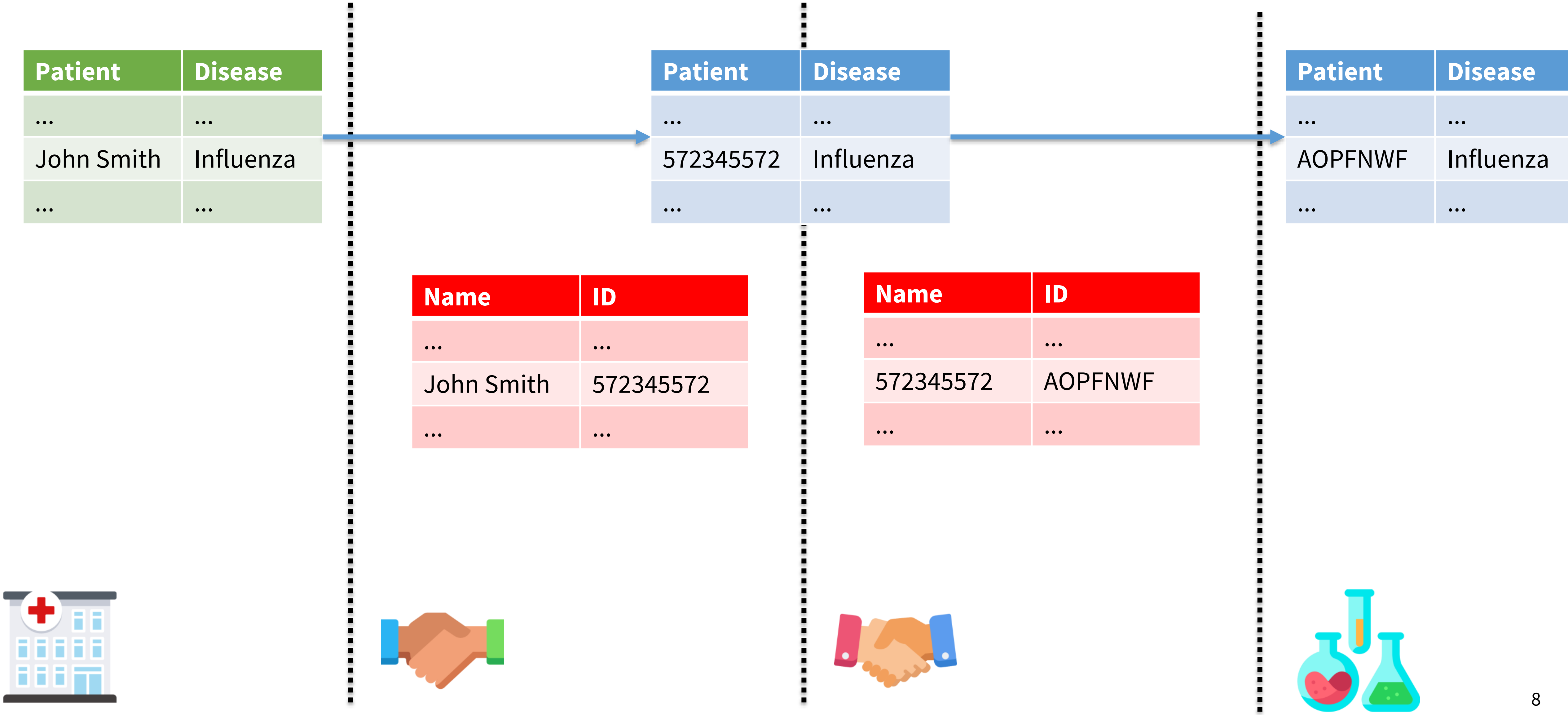
Patient	Disease	Age
…	…	…
John Smith	Influenza	57
…	…	…

Name	ID
…	…
John Smith	572345572
…	…

Patient	Disease	Age
…	…	…
572345572	Influenza	57
…	…	…

Icons made by Freepik from www.flaticon.com

Pseudonymisation Architectures (3)



Patient	Disease
…	…
John Smith	Influenza
…	…

Patient	Disease
…	…
572345572	Influenza
…	…

Patient	Disease
…	…
AOPFNWF	Influenza
…	…

Name	ID
…	…
John Smith	572345572
…	…

Name	ID
…	…
572345572	AOPFNWF
…	…

Icons made by Freepik from www.flaticon.com

One-time use – No re-identification required

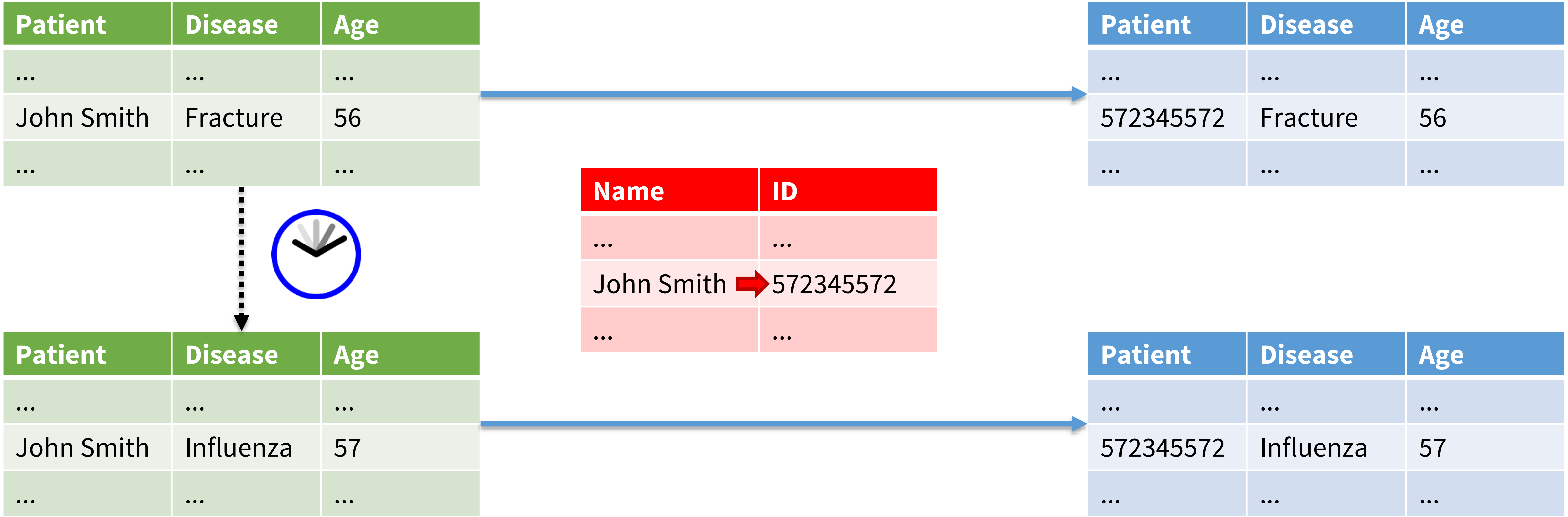
- Example usage: statistical analysis
- No mapping between patient name and patient ID required
- Can use random patient ID (or even remove the patient ID)
- No mapping information must be stored
- No pseudonymisation → anonymisation

Patient	Disease	Age
…	…	…
John Smith	Influenza	57
…	…	…

Patient	Disease	Age
…	…	…
572345572	Influenza	57
…	…	…

Follow-up use – No re-identification required

- Follow-up participations shall be linkable



Patient	Disease	Age
…	…	…
John Smith	Fracture	56
…	…	…

Patient	Disease	Age
…	…	…
572345572	Fracture	56
…	…	…

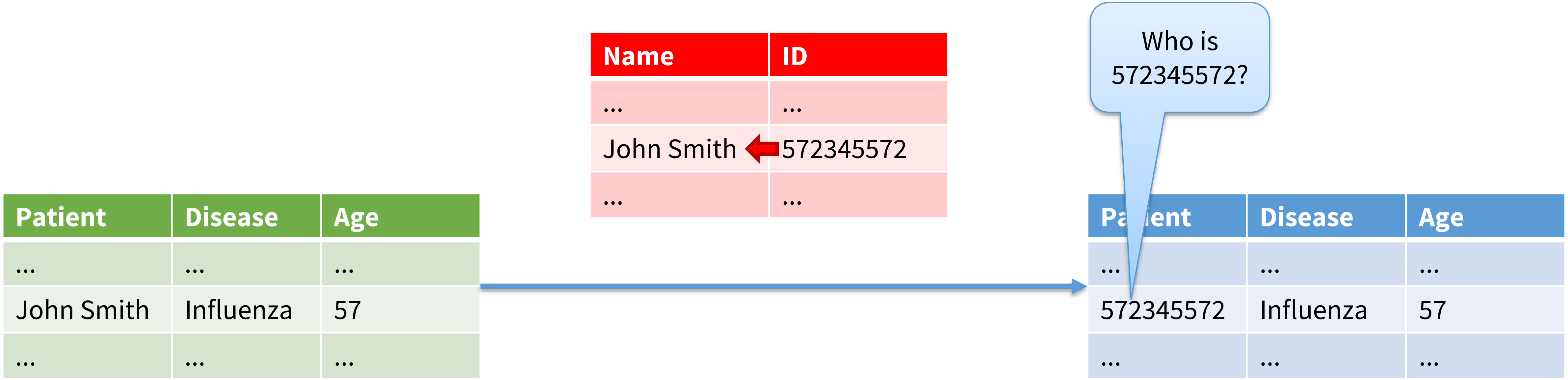
Name	ID
…	…
John Smith →	572345572
…	…

Patient	Disease	Age
…	…	…
John Smith	Influenza	57
…	…	…

Patient	Disease	Age
…	…	…
572345572	Influenza	57
…	…	…

One-time use – Re-identification possible

- Patients shall be identifiable to contact them
- Used only in case of “emergency”, e.g. serious disease found after analysis of data or blood/tissue sample



Name	ID
…	…
John Smith	572345572
…	…

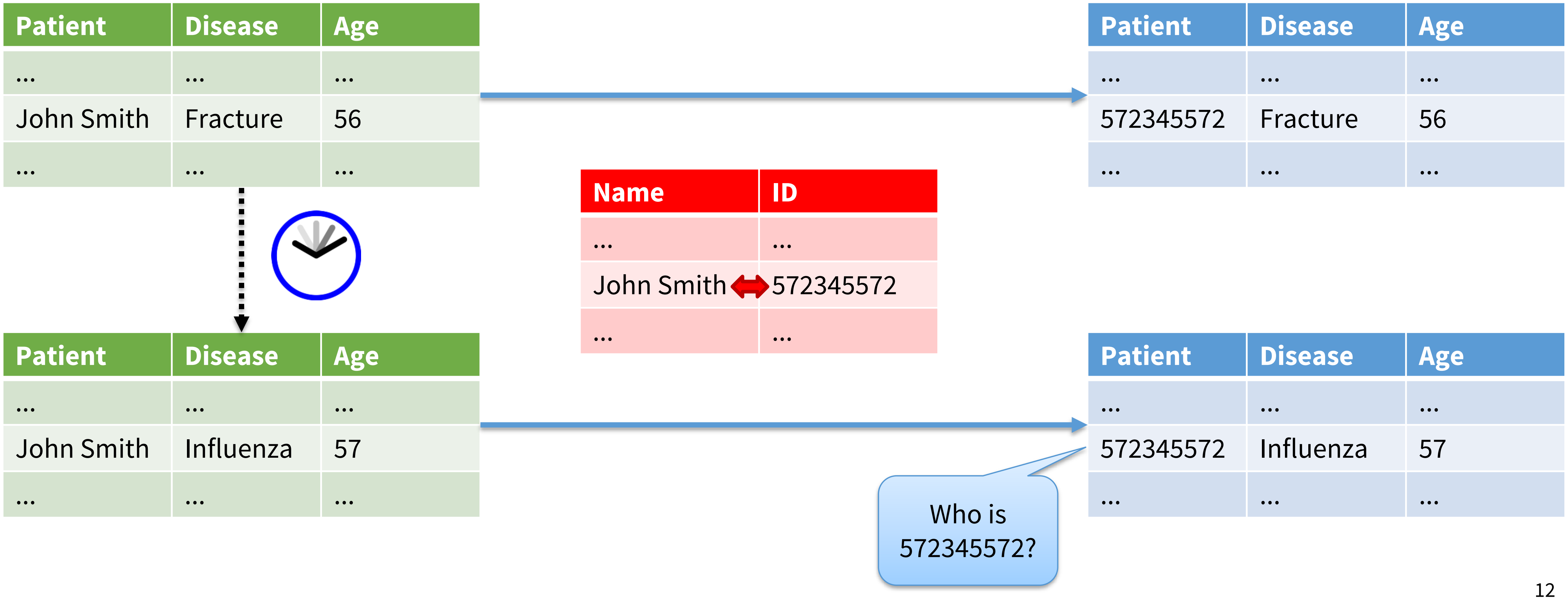
Patient	Disease	Age
…	…	…
John Smith	Influenza	57
…	…	…

Who is 572345572?

Patient	Disease	Age
…	…	…
572345572	Influenza	57
…	…	…

Follow-up use – Re-identification possible

- Follow-up visits shall be linkable + patients shall be identifiable



Patient	Disease	Age
…	…	…
John Smith	Fracture	56
…	…	…

Patient	Disease	Age
…	…	…
John Smith	Influenza	57
…	…	…

Name	ID
…	…
John Smith ⟷	572345572
…	…

Patient	Disease	Age
…	…	…
572345572	Fracture	56
…	…	…

Patient	Disease	Age
…	…	…
572345572	Influenza	57
…	…	…

Who is 572345572?

Implementation of Pseudonymisation (1)

- Random number
  - Mapping table between name and pseudonym must be explicitly stored
  - Table allows resolution in both directions
- Counter
  - Chronology of pseudonyms might reveal identities
  - (otherwise: same properties as random number)
- Symmetric encryption
  - Pseudonym is cipher text of name
  - Only storage of secret key is required
  - Allows resolution in both direction (encryption / decryption)

Name	ID
Mary Doe	173965429
John Smith	572345572
Adam James	938426463
…	…

Name	ID
Mary Doe	783956301
John Smith	783956302
Adam James	783956303
…	…

Implementation of Pseudonymisation (2)

- Cryptographic hash function
  - Pseudonym is hash value of name
  - No storage required
  - Backwards resolution not directly possible (just with huge effort)
  - Insecure (see next slide)
- Cryptographic hash function + secret (e.g. HMAC)
  - Pseudonym is hash value of name + secret value
  - Just storage of secret required
  - Backwards resolution not directly possible (just with huge effort)

Attacks on Pseudonymity

- Insider:
  - Knows mapping table / secrets (or can easily gain this)
- Outsider:
  - In case of hash-based pseudonym:
    - guess hash function
    - generate (valid) pseudonyms
    - perform pre-image attacks (brute-force, dictionary, rainbow-table)
  - Gain insider knowledge / access

Attacks on Anonymity

- Quasi-identifier can reveal the identity

Patient	Age	Gender	City	Disease
173965429	57	female	Hamburg	Gastric ulcer
572345572	18	male	Amsterdam	Influenza
938426463	92	female	Süderbrarup	Fracture of leg
824678234	37	male	Brussles	Herpes zoster

- Example countermeasure: generalization

Patient	Age	Gender	Region	Disease
173965429	> 50	female	Germany	Gastric ulcer
572345572	≤ 50	male	Benelux	Influenza
938426463	> 50	female	Germany	Fracture of leg
824678234	≤ 50	male	Benelux	Herpes zoster

Attacks on Anonymity

- Homogeneity Attack

Patient	Age	Gender	Region	Disease
173965429	> 50	female	Germany	Gastric ulcer
572345572	≤ 50	male	Benelux	Influenza
938426463	> 50	female	Germany	Gastric ulcer
824678234	≤ 50	male	Benelux	Herpes zoster

- Background Knowledge Attack

Patient	Age	Gender	Region	Disease
173965429	> 50	female	Germany	Gastric ulcer
572345572	≤ 50	male	Benelux	Fracture of leg
938426463	> 50	female	Germany	Herpes zoster
824678234	≤ 50	male	Benelux	Appendix

Bart likes snow boarding and has lost his appendix 5 years ago

Thank you for your attention!