Die wunderbare Welt kryptographischer Pseudonyme

Europäischer Datenschutztag 2026

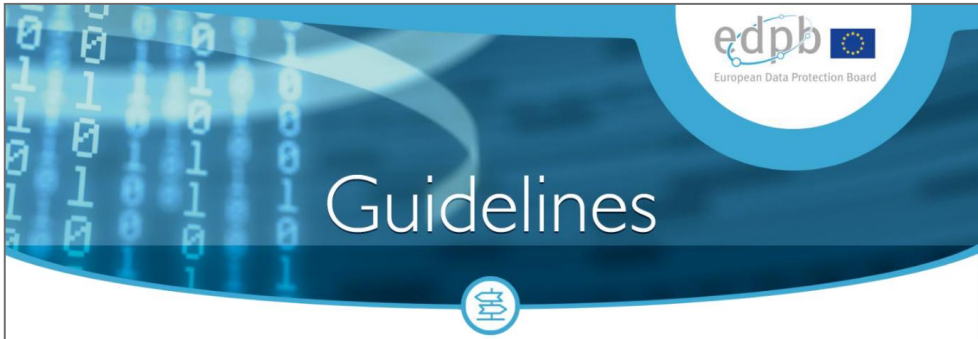
Prof. Dr. Anja Lehmann

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Pseudonymization

Pseudonymisation is defined in Art. 4(5) GDPR as "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."*

Disclaimer
This talks: pseudonymization from the perspective of cryptographic research



Pseudonymization

- Pseudonymization = replacing direct identifier with random-looking substitute

Name	Credit	IBAN
Alice Doe	52.650	CH56 0483 5084 1385 0100 0

Nym	Credit
xH2ban6	52.650

Pseudonymization

- Pseudonymization = replacing direct identifier with random-looking substitute

Name	Credit	IBAN
Alice Doe	52.650	CH56 0483 5084 1385 0100 0

Nym	Credit
xH2ban6	52.650

- Requirements: pseudonymization must be
  - Secure, i.e., infeasible to determine **uid** from **nym**
  - Consistent, i.e., referential integrity must be preserved
  - Dynamic, i.e., data can be pseudonymized on the fly

UserID	Credit
Alice	€ 8.000
Bob	€ 3.500

UserID	Credit
Alice	€ 599

Nym	Credit
xH2ban6	€ 8.000
P3b0Ws	€ 3.500

Nym	Credit
xH2ban6	€ 599

Pseudonymization | Current Solutions

87. Two classes of replacement procedures are commonly applied as pseudonymising transformations: cryptographic algorithms and lookup tables.[22]

UserID	Nym
Alice	xH2ban6
Bob	P89dyCk
Carol	912uj4V

Pseudonymization | Current Solutions

87. Two classes of replacement procedures are commonly applied as pseudonymising transformations: cryptographic algorithms and lookup tables.[22]

UserID	Nym
Alice	xH2ban6
Bob	P89dyCk
Carol	912uj4V

- Deterministic one-way functions, e.g. Hash function $H(uid) \rightarrow nym$

Pseudonymization | Current Solutions

87. Two classes of replacement procedures are commonly applied as pseudonymising transformations: cryptographic algorithms and lookup tables.[22]

UserID	Nym
Alice	xH2ban6
Bob	P89dyCk
Carol	912uj4V

- Deterministic one-way functions, e.g. Hash function $H(uid) \rightarrow nym$

- But: deterministic functions are vulnerable to offline attacks:

$$H(Bob) = nym\,?, H(Eve) = nym\,?, \ldots$$

Pseudonymization | Current Solutions

87. Two classes of replacement procedures are commonly applied as pseudonymising transformations: cryptographic algorithms and lookup tables.[22]

UserID	Nym
Alice	xH2ban6
Bob	P89dyCk
Carol	912uj4V

- Deterministic one-way functions, e.g. Hash function $H(uid) \rightarrow nym$

- But: deterministic functions are vulnerable to offline attacks:

$$H(Bob) = nym\,?, H(Eve) = nym\,?, \ldots$$

- Pseudonymization must be based on strong secret key, e.g., $PRF(key, uid) \rightarrow nym$

Pseudonymization | Current Solutions

87. Two classes of replacement procedures are commonly applied as pseudonymising transformations: cryptographic algorithms and lookup tables.[22]

UserID	Nym
Alice	xH2ban6
Bob	P89dyCk
Carol	912uj4V

- Deterministic one-way functions, e.g. Hash function $H(uid) \rightarrow nym$

- But: deterministic functions are vulnerable to offline attacks:

$$H(Bob) = nym\,?, H(Eve) = nym\,?, \ldots$$

- Pseudonymization must be based on strong secret key, e.g., $PRF(key, uid) \rightarrow nym$

In theory: „Let K be a key“ In practice: Well, its not that easy ....

Challenge 1: Distributed Pseudonymization & Consistency

- Pseudonymization from multiple sources while preserving consistency

UserID	X-Ray
Alice	Picture-1
Bob	Picture-12

UserID	Doctor A
Alice	Measurement

UserID	Doctor B
Alice	Eval

Nym	X-Ray
xH2ban6	Picture-1
P3b0Ws	Picture-12

Nym	Doctor A
xH2ban6	Measurement

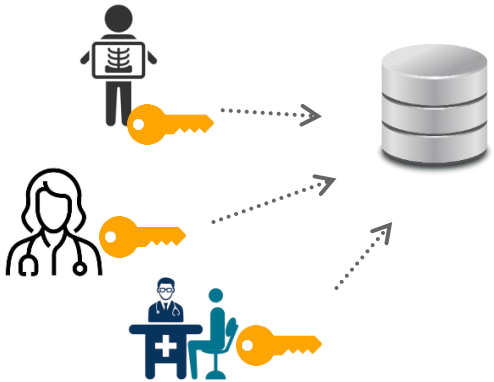
Nym	Doctor B
xH2ban6	Eval

How to ensure consistent and secure pseudonymization when data is pushed by many, distributed entities ?

Distributed Pseudonymization | Existing Solutions

- Secure pseudonymization requires strong cryptographic keys or secret lookup table

Key is shared with all entities

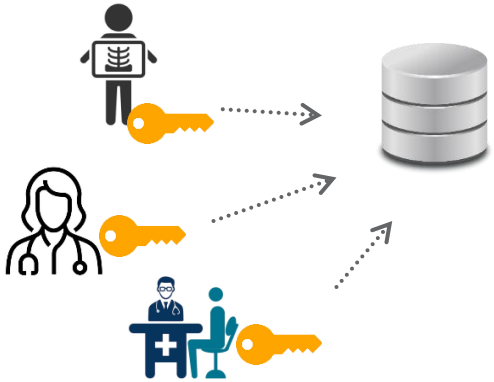


– Replication of keys reduces security

Distributed Pseudonymization | Existing Solutions

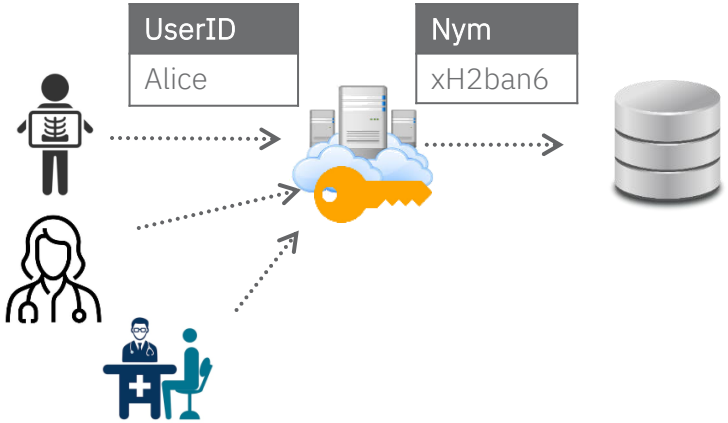
- Secure pseudonymization requires strong cryptographic keys or secret lookup table

Key is shared with all entities



– Replication of keys reduces security

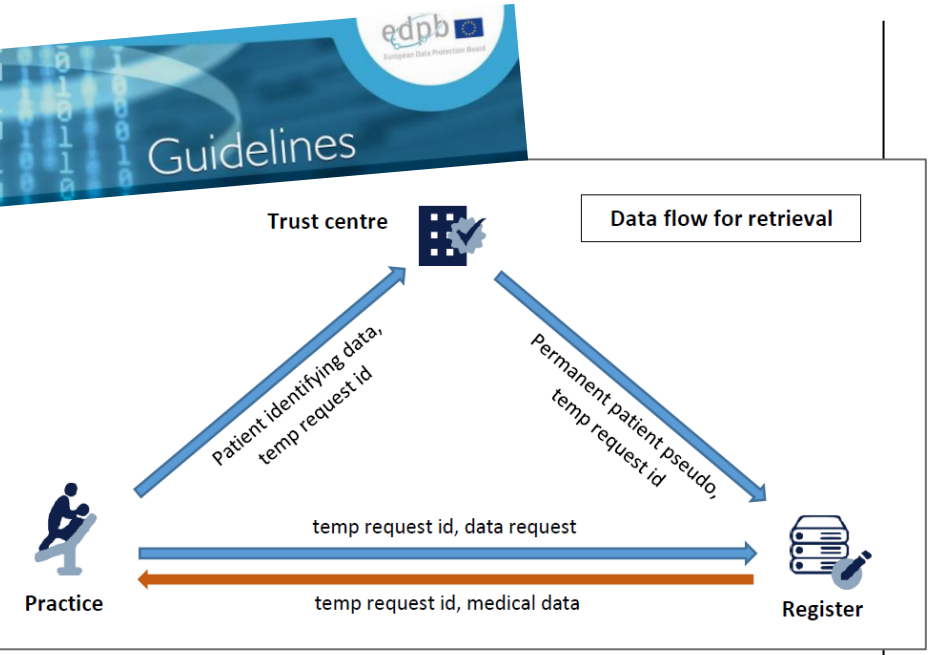
Central trusted service (TTP)



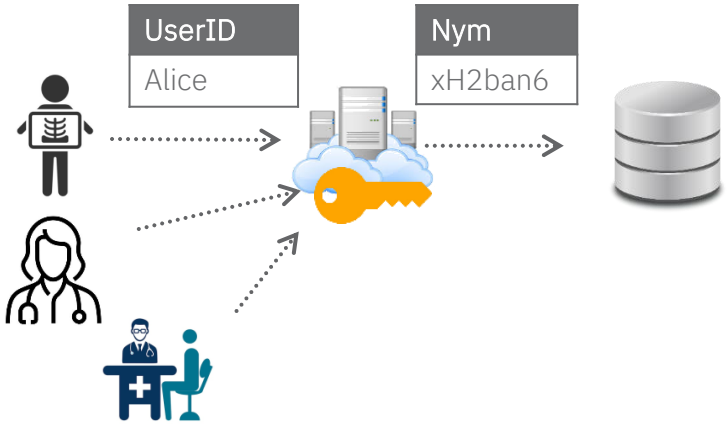
+ Simple and secure deployment

Distributed Pseudonymization | Existing Solutions

- Secure pseudonymization requires strong cryptographic keys or secret lookup table



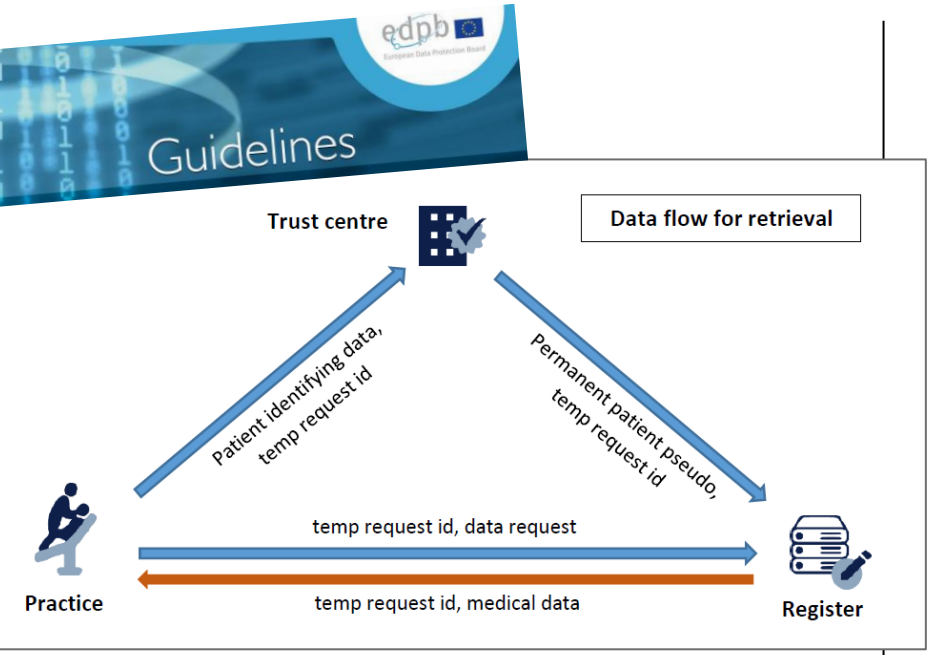
Central trusted service (TTP)



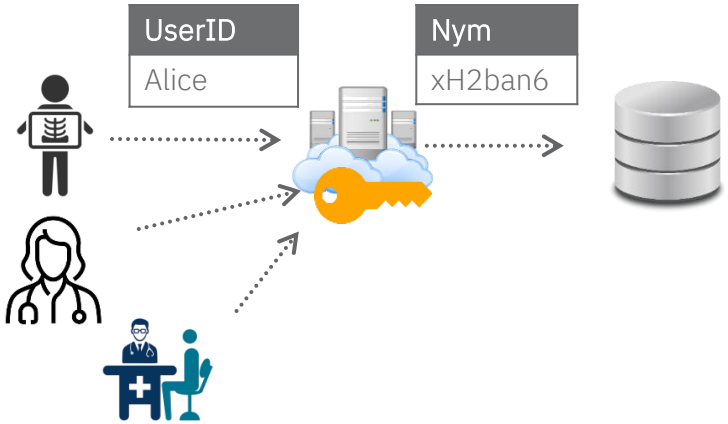
\+ Simple and secure deployment

Distributed Pseudonymization | Existing Solutions

- Secure pseudonymization requires strong cryptographic keys or secret lookup table



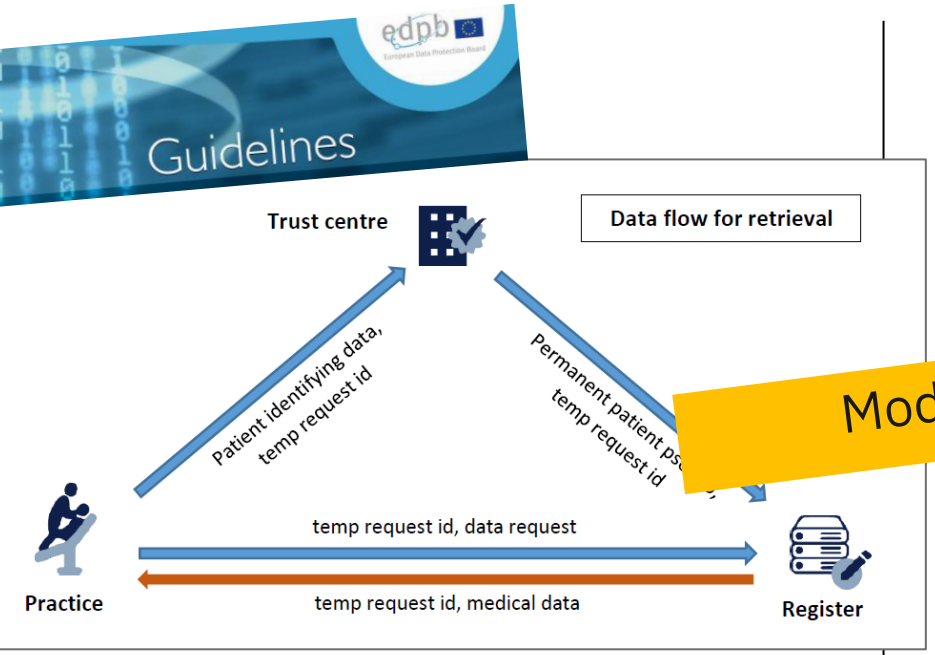
Central trusted service (TTP)



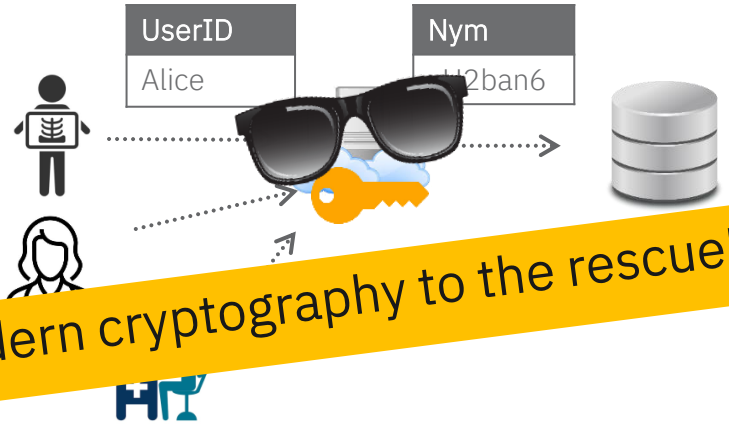
+ Simple and secure deployment
– TTP learns all metadata
– TTP must be fully trusted → privacy risk itself!

Distributed Pseudonymization | Existing Solutions

- Secure pseudonymization requires strong cryptographic keys or secret lookup table



Central trusted service (TTP)



Modern cryptography to the rescue!

\+ Simple and secure deployment

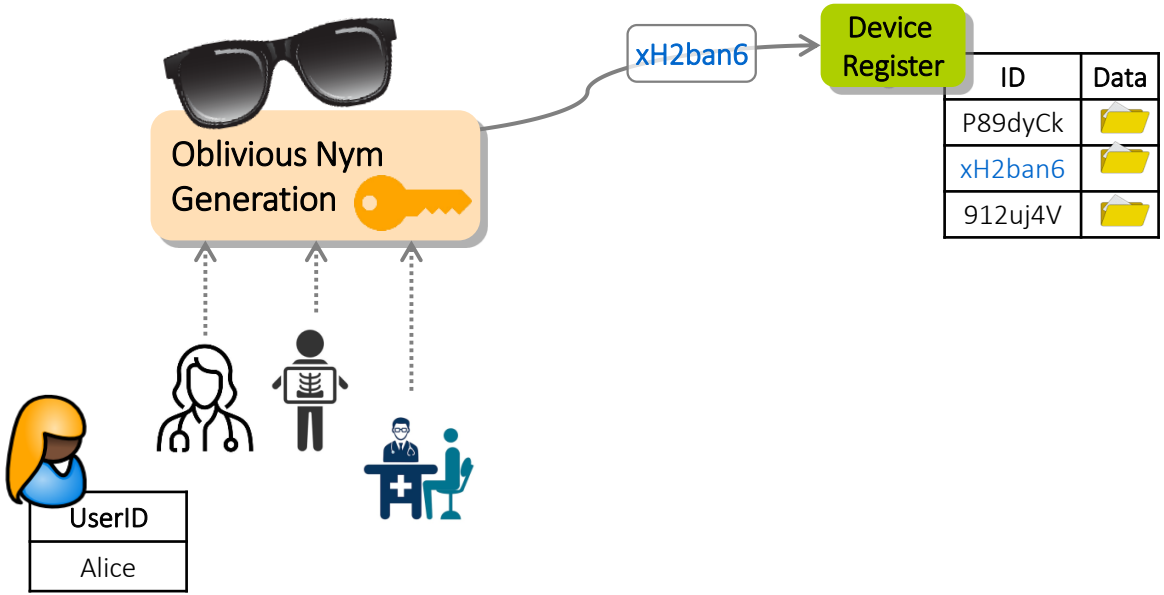
– ~~TTP learns all metadata~~

– ~~TTP must be fully trusted → privacy risk itself!~~

Distributed Pseudonymization | *Oblivious* Pseudonymization

- Central entity **blindly** computes consistent pseudonyms $nym = PRF(key, uid)$

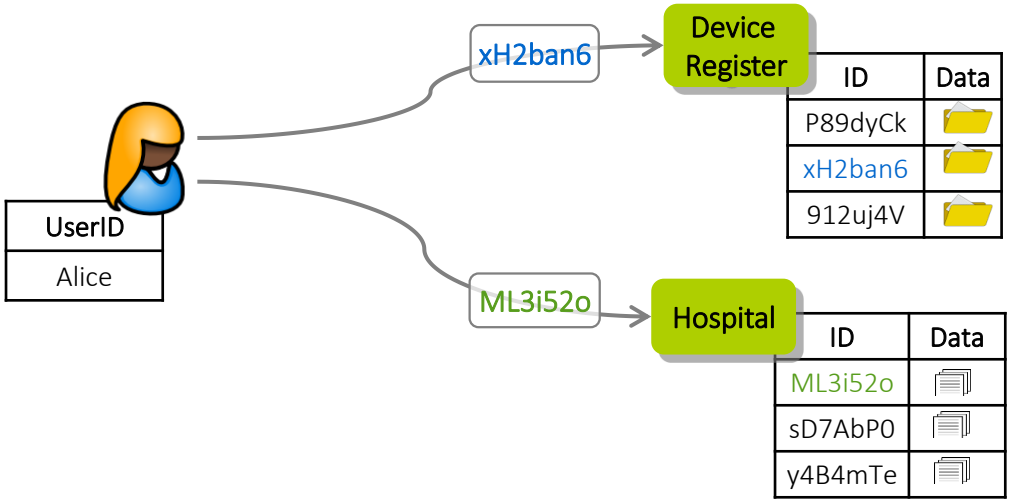
  → learns nothing about uid or nym!



[L19] *ScrambleDB: Oblivious (Chameleon) Pseudonymization-as-a-Service.* Lehmann. PETS 2019

Restrict Consistency to Domain Pseudonyms

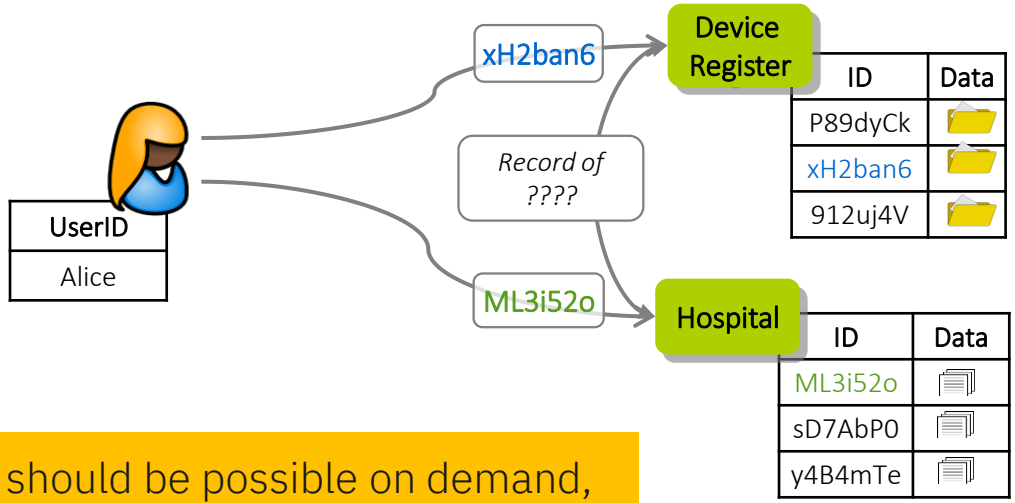
- Limit consistency of pseudonyms to minimally scoped domains
- User has domain-specific pseudonyms → unlinkable across domains



Challenge 2: Data Exchange with Unlinkable Pseudonyms

- How to exchange and correlate pseudonymous data ?

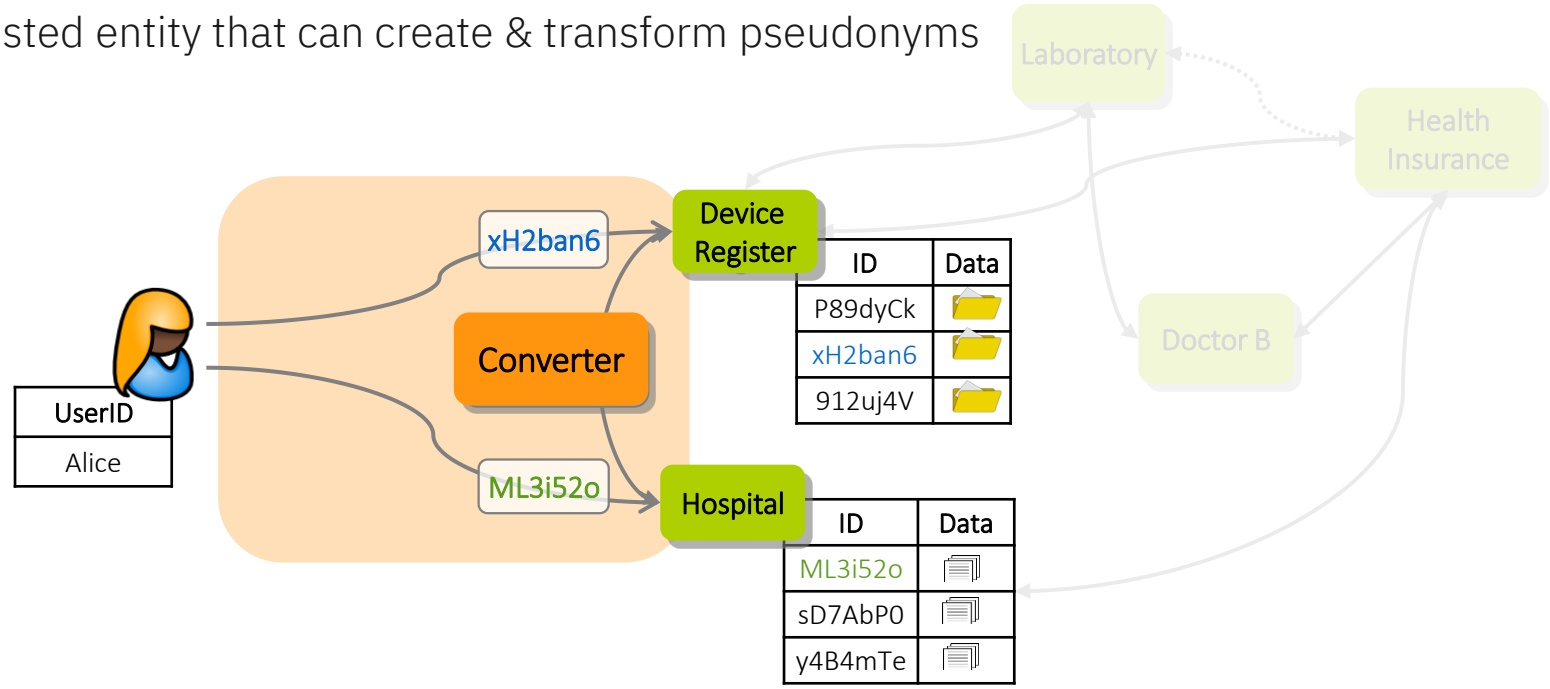
  E.g., eHealth records, social security system



Correlation should be possible on demand, but in a controlled way

Linkable Pseudonyms? *Trusted* Converter

- Central trusted entity that can create & transform pseudonyms

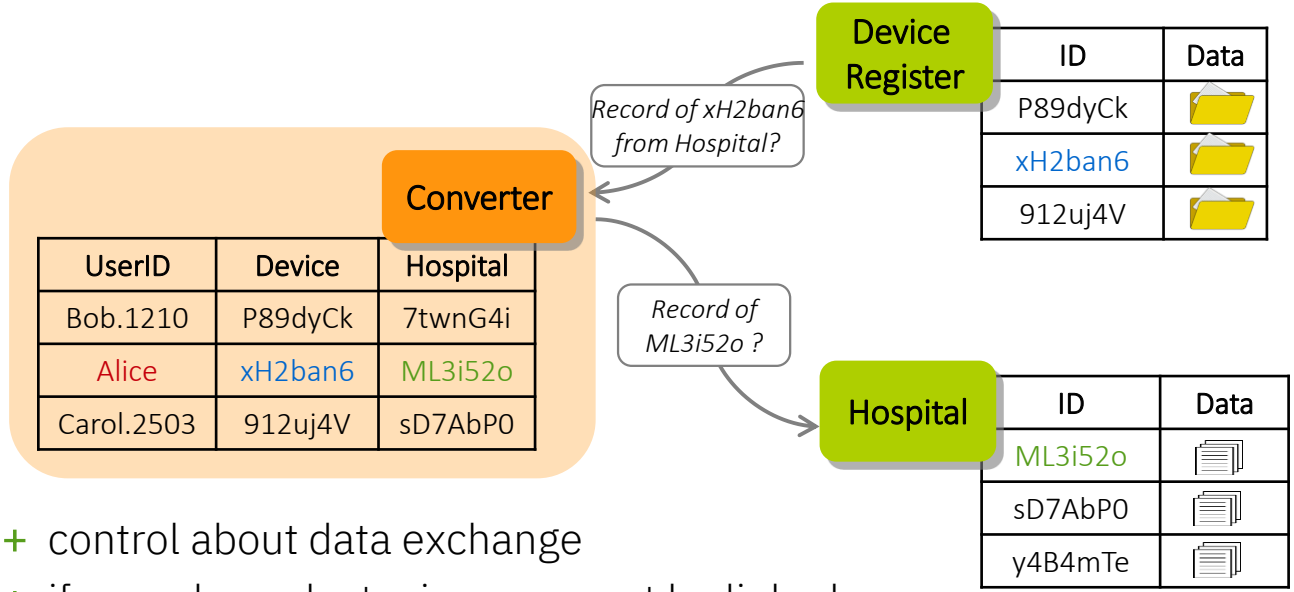


edpb European Data Protection Board

123. Several approaches to implementing pseudonymising transformations allow for the controlled linking of pseudonymised data held by different controllers: Sharing of pseudonymisation secrets among controllers, jointly using a trusted service provider for carrying out the pseudonymisation,

Linkable Pseudonyms? *Trusted* Converter

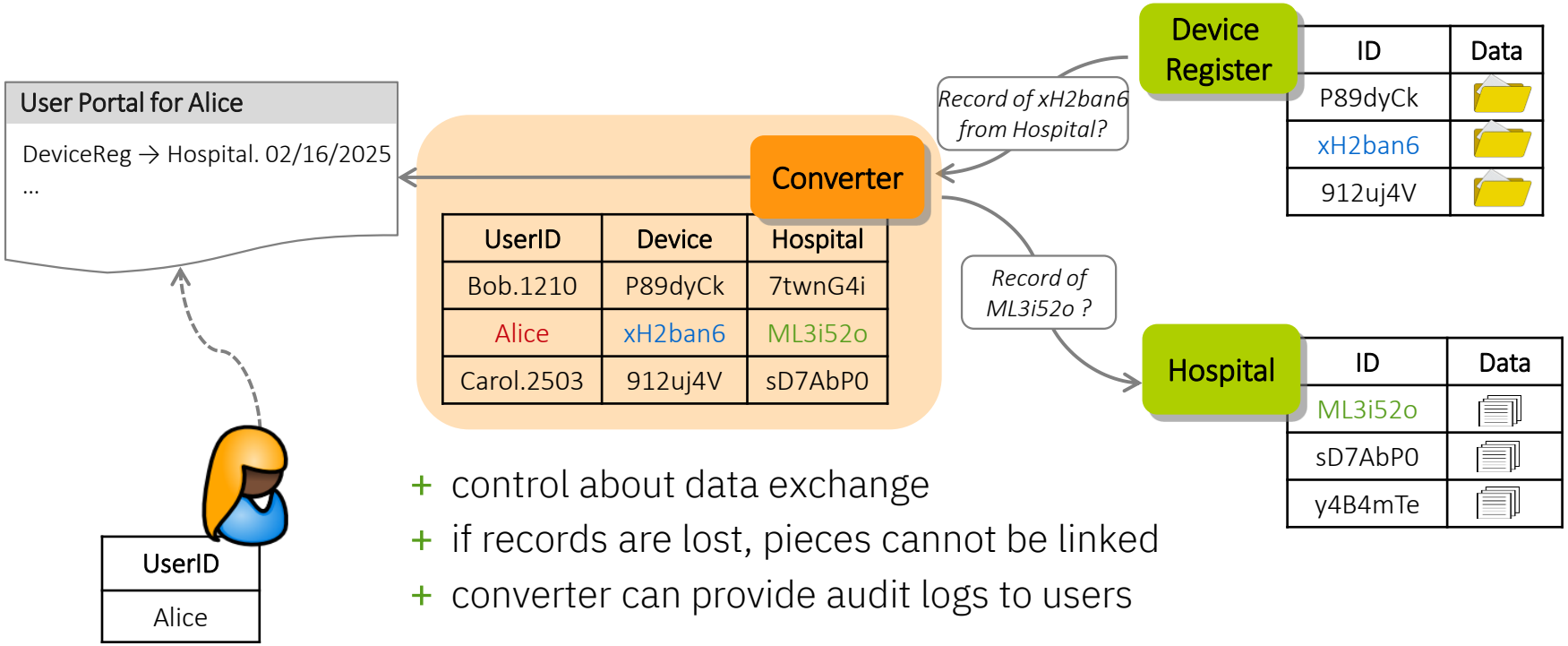
- Central trusted entity that can create & transform pseudonyms



+ control about data exchange
+ if records are lost, pieces cannot be linked

Linkable Pseudonyms? *Trusted* Converter

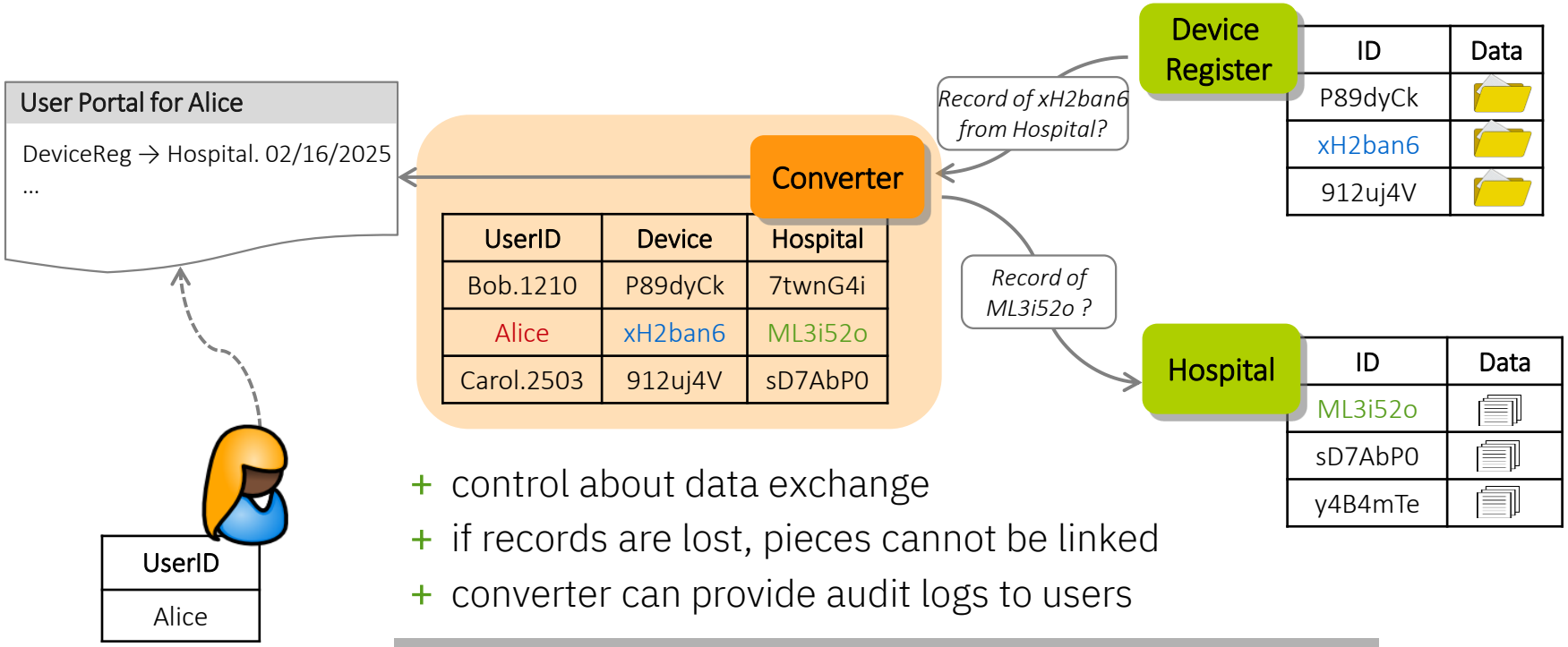
- Central trusted entity that can create & transform pseudonyms



+ control about data exchange
+ if records are lost, pieces cannot be linked
+ converter can provide audit logs to users

Linkable Pseudonyms? *Trusted* Converter

- Central trusted entity that can create & transform pseudonyms



\+ control about data exchange

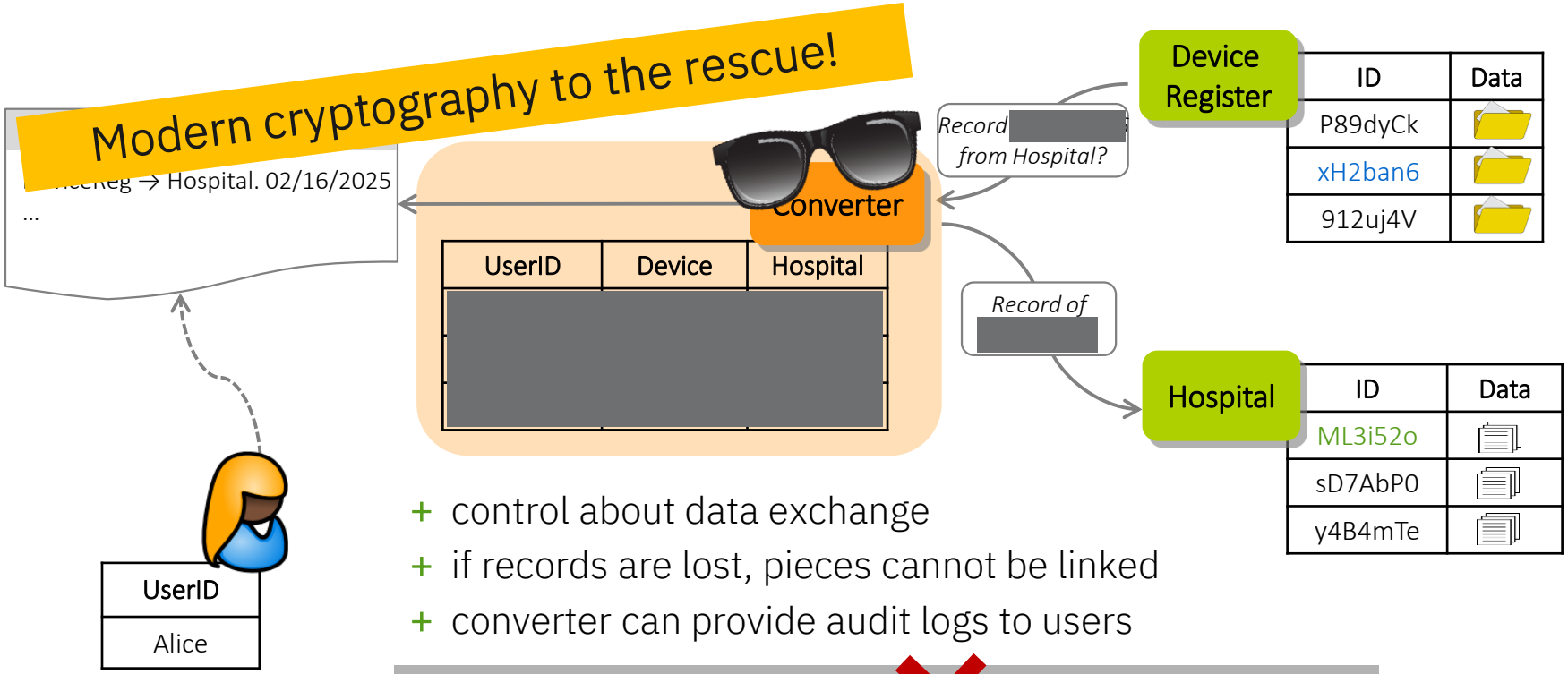
\+ if records are lost, pieces cannot be linked

\+ converter can provide audit logs to users

– converter learns all requests & knows all correlations

Linkable Pseudonyms? *Oblivious* Converter

- Central trusted entity that can create & transform pseudonyms ***blindly***

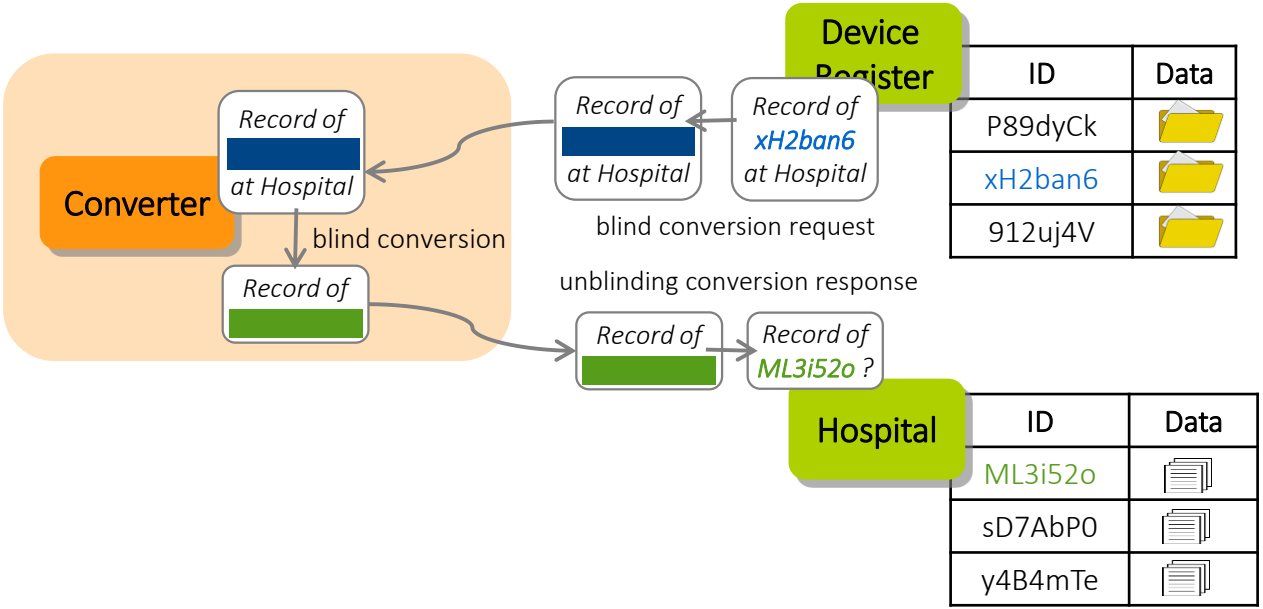


+ control about data exchange
+ if records are lost, pieces cannot be linked
+ converter can provide audit logs to users

– converter learns all requests & knows all correlations

Blindly Convertible Pseudonyms | Pseudonym Conversion

- Only converter can link & transform pseudonyms, but does so in a blind way



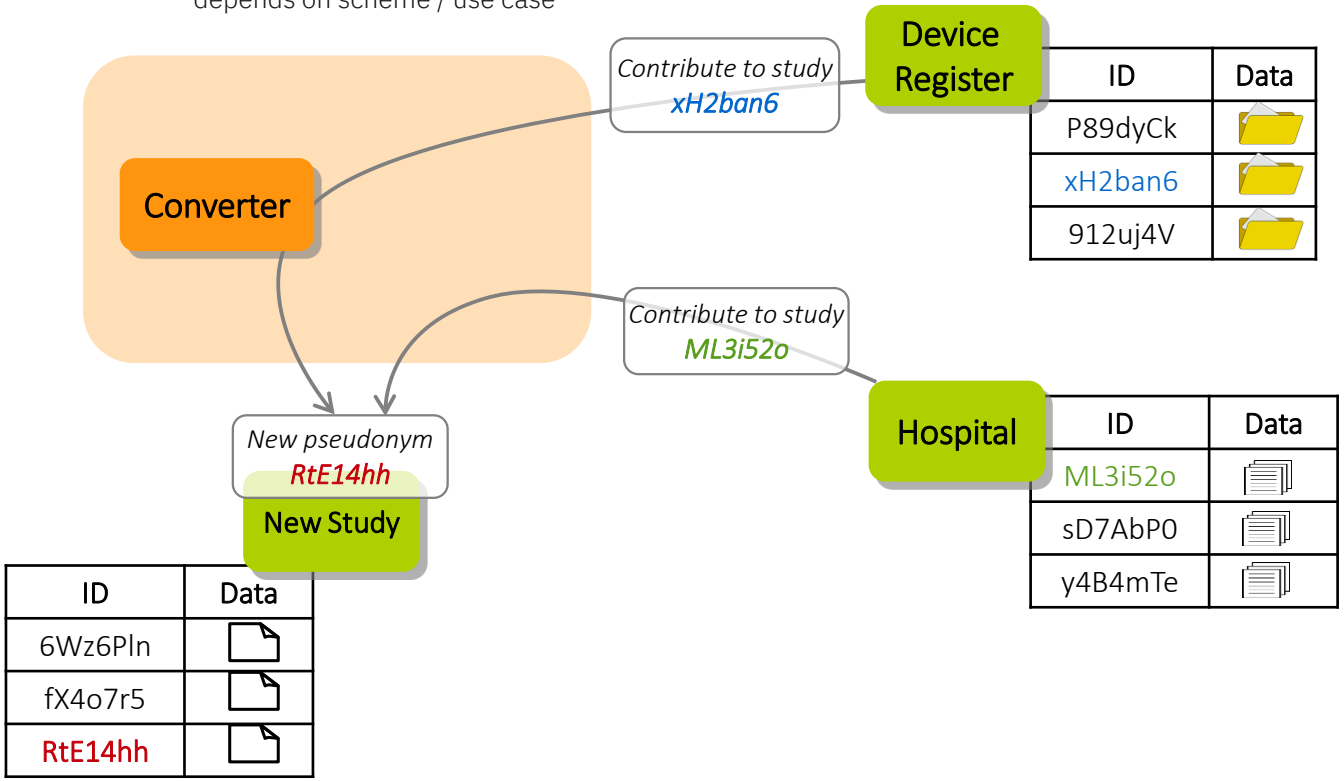
[CL15] *Privacy for Distributed Databases via (Un)linkable Pseudonyms*. Camenisch, Lehmann. ACM CCS 2015.

[VJM+16] *Polymorphic encryption and pseudonymisation for personalised healthcare.* Verheul et al., Technical Report 2016.

Blindly Convertible Pseudonyms | Consistent Joins

- Pseudonym conversions are transitive*, unlinkable data can be aggregated

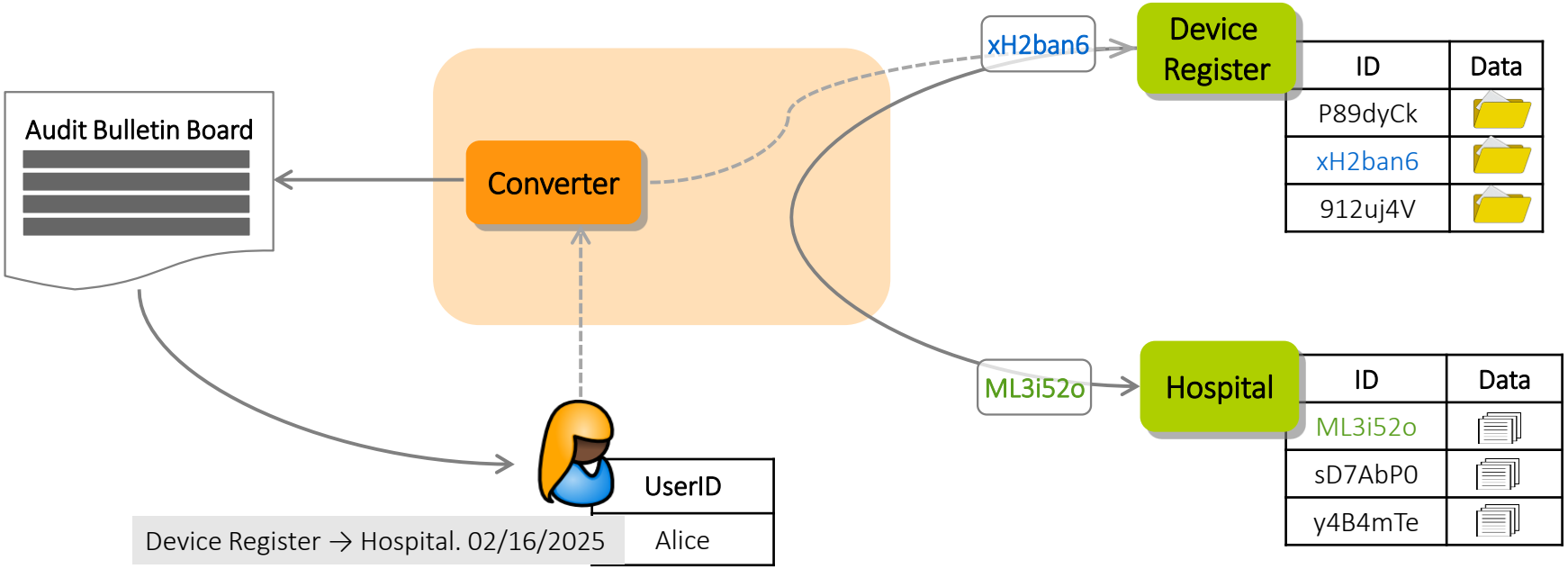
*depends on scheme / use case



[LMS26] *Multi-Party Private Join*. Lehmann, Mouchet, Sidorenko. PETS 2026

Blindly Convertible Pseudonyms | User Audits

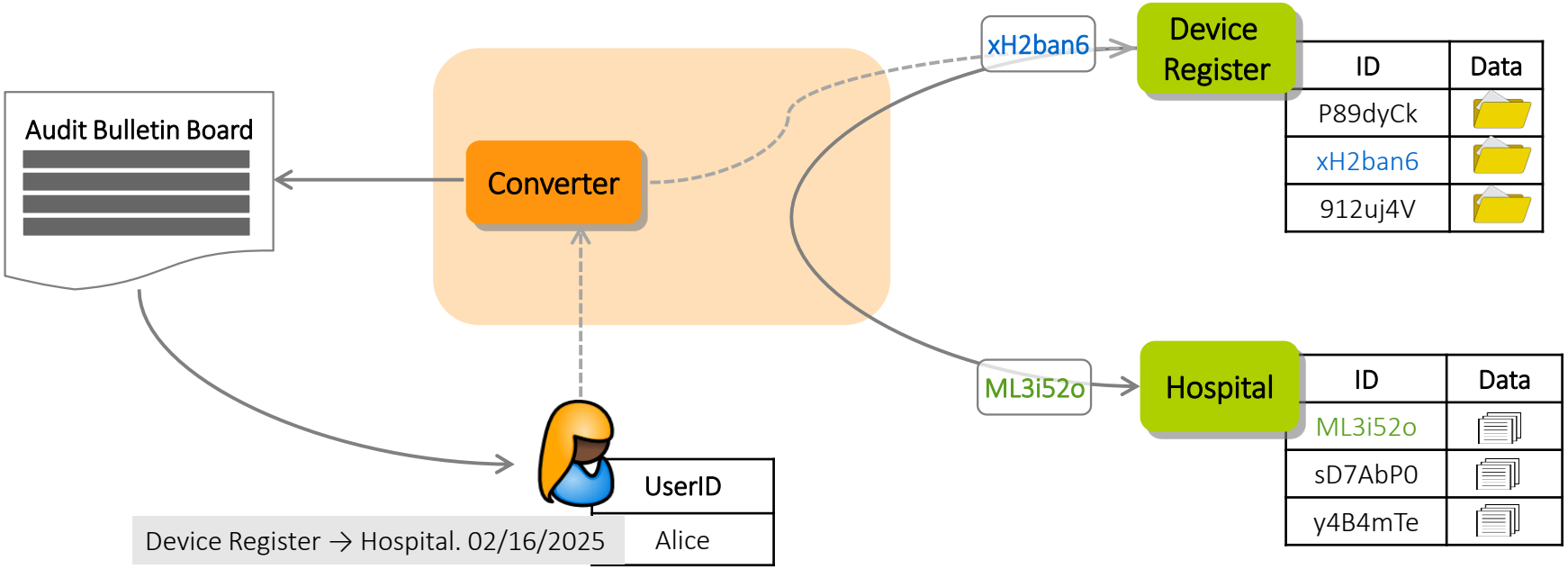
- Every pseudonym conversion triggers blind generation of audit log entry



[CL17] *Privacy-Preserving User-Auditable Pseudonym Systems.* Camenisch, Lehmann. IEEE EuroS&P 2017

Blindly Convertible Pseudonyms | User Audits

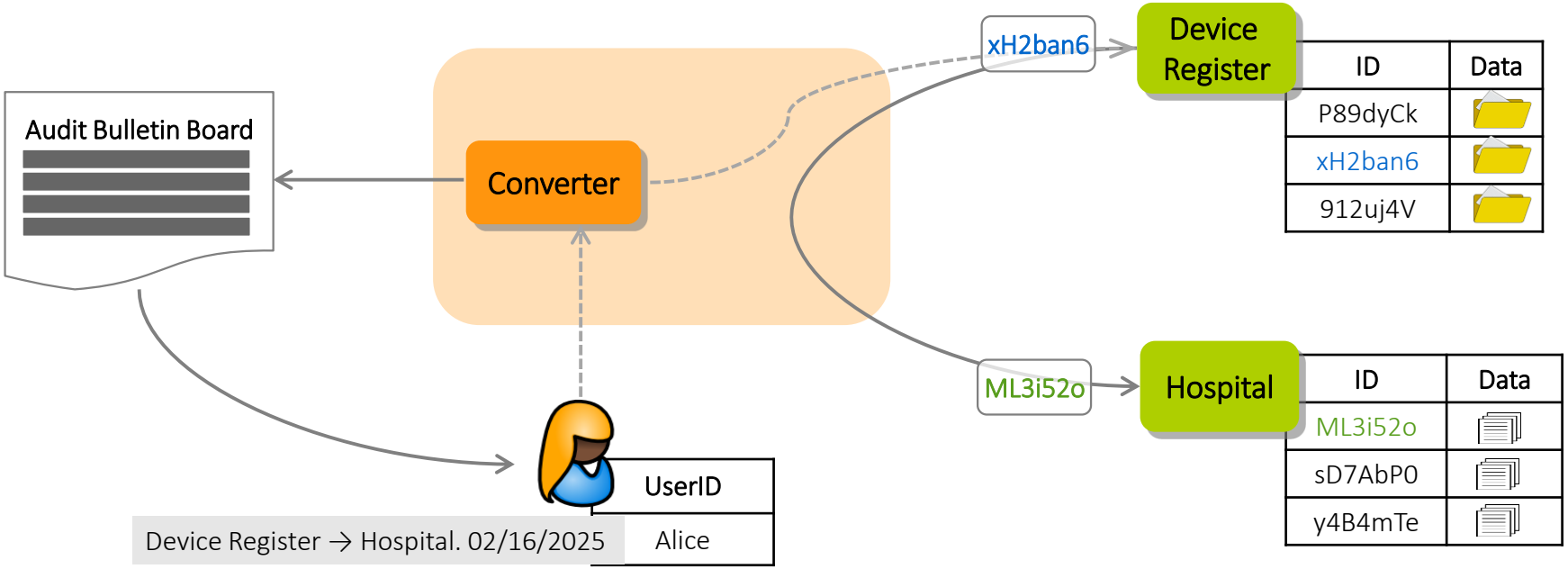
- Every pseudonym conversion triggers blind generation of audit log entry



(Convertible) pseudonyms rely on disclosure of unique identifiers

Blindly Convertible Pseudonyms | User Audits

- Every pseudonym conversion triggers blind generation of audit log entry

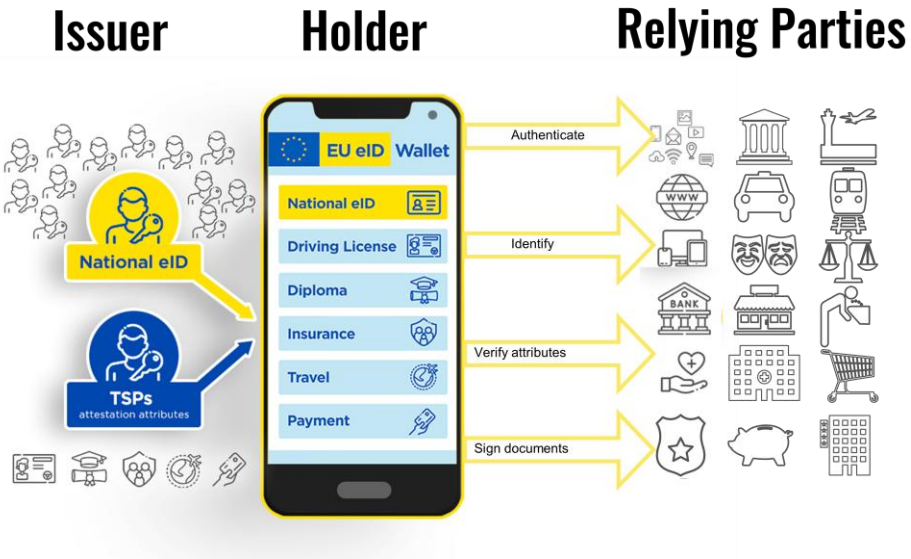


(Convertible) pseudonyms rely on disclosure of unique identifiers

… or on a *pseudonymous* digital identity → EUDI ?

European Digital Identity Wallet (EUDI)

- All EU member states must provide a EUDI Wallet by end of 2026

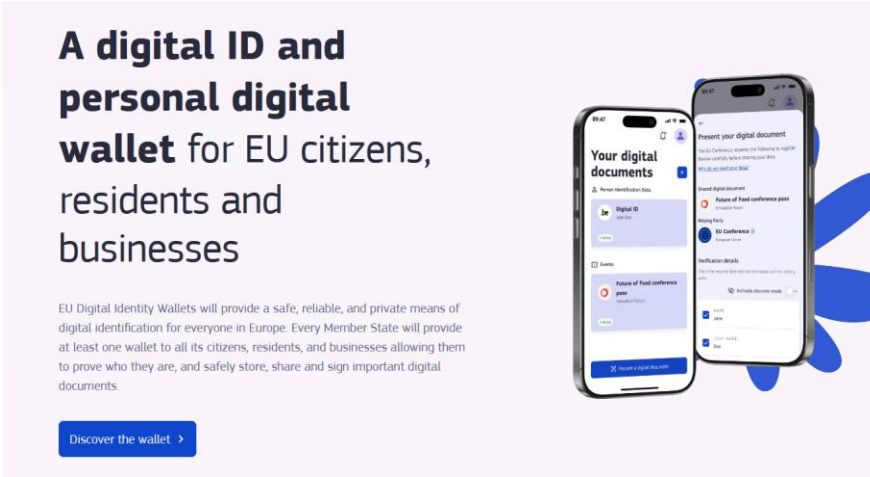


"fully mobile, secure and user-friendly"

https://ec.europa.eu/digital-building-blocks/sites/display/EUDIGITALIDENTITYWALLET/EU+Digital+Identity+Wallet+Home

European Digital Identity Wallet (EUDI)

- All EU member states must provide a EUDI Wallet by end of 2026



"fully mobile, secure and user-friendly"

https://ec.europa.eu/digital-building-blocks/sites/display/EUDIGITALIDENTITYWALLET/EU+Digital+Identity+Wallet+Home

EUDI | eIDAS Regulation → Wallet in 2026

- EU regulation eIDAS 2.0 entered into force in May 2024
- Regulation mandates strong privacy & security properties

https://eur-lex.europa.eu/eli/reg/2024/1183/oj

"securely [..] authenticate to relying parties [..] while ensuring **selective disclosure of data** [..]

enable privacy-preserving techniques which ensure **unlinkability** [..]

possibility of users to access services through the use of **pseudonyms** [..]

providers should ensure **unobservability** by not collecting data and not having insight into the transactions of the users [..]

EUDI | eIDAS Regulation → Wallet in 2026

- EU regulation eIDAS 2.0 entered into force in May 2024
- Regulation mandates strong privacy & security properties

https://eur-lex.europa.eu/eli/reg/2024/1183/oj

"securely [..] authenticate to relying parties [..] while ensuring **selective disclosure of data** [..]

enable privacy-preserving techniques which ensure **unlinkability** [..]

possibility of users to access services through the use of **pseudonyms** [..]

providers should ensure **unobservability** by not collecting data and not having insight into the transactions of the users [..]

"European Digital Identity Wallets should include a functionality to generate **user-chosen and managed pseudonyms**, to authenticate when accessing online services. [..]

right of the users **to use freely chosen pseudonyms**. [..]

shall enable the user [..] to: **generate pseudonyms and store them encrypted** and locally within the European Digital Identity Wallet;

[..] the **use of pseudonyms that are chosen by the user shall not be prohibited**"

EUDI | eIDAS Regulation → Wallet in 2026

- EU regulation eIDAS 2.0 entered into force in May 2024
- Regulation mandates strong privacy & security properties

https://eur-lex.europa.eu/eli/reg/2024/1183/oj

"securely [..] authenticate to relying parties [..] while ensuring **selective disclosure of data** [..]

enable privacy-preserving techniques which ensure **unlinkability** [..]

possibility of users to access services through the use of **pseudonyms** [..]

providers should ensure **unobservability** by not [illegible] insight into the transacti

Pseudonym = Donald Duck ??

"European Digital Identity Wallets should include a functionality to generate **user-chosen and managed pseudonyms**, to authenticate when accessing online services. [..]

right of the users **to use freely chosen pseudonyms**. [..]

shall enable the user [..] to: **generate pseudonyms and store them encrypted** and locally within the European Digital Identity Wallet;

[..] the **use of pseudonyms that are chosen by the user shall not be prohibited**"

EUDI | Potential for Pseudonymous Authentication

Know-Your-Customer
open bank account

Age Proofs

EUDI | Potential for Pseudonymous Authentication

Know-Your-Customer
open bank account

This is where **pseudonymous authentication** can have biggest impact

Age Proofs

EUDI | Potential for Pseudonymous Authentication

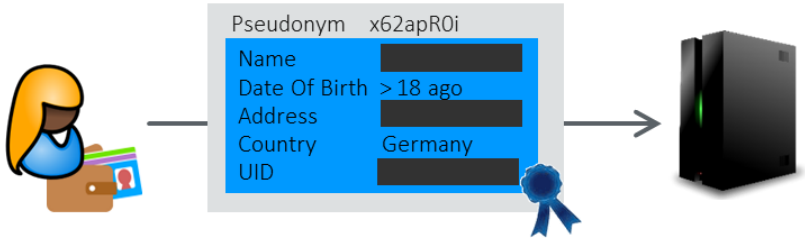
Security

Know-Your-Customer
open bank account

This is where **pseudonymous authentication** can have biggest impact

Privacy

Age Proofs



- Pseudonym derived from certified (but secret) UID: $nym = F(uid)$
  Establishes verifiable and unforgeable pseudonymous identity
  Pseudonymous authentication is unlinkable (with ZKPs)

[MLS25] *A Brief Note on Cryptographic Pseudonyms for Anonymous Credentials.* Mayrhofer, Lehmann, shelat

EUDI | Potential for Pseudonymous Authentication

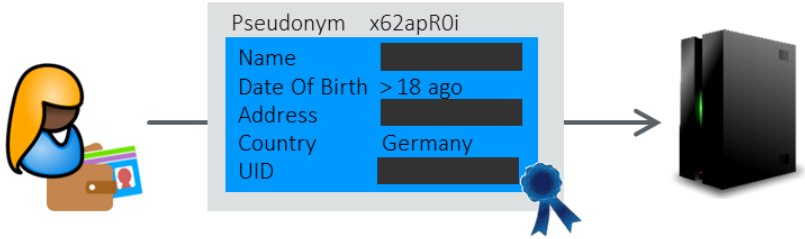
Security

Know-Your-Customer
open bank account

This is where **pseudonymous authentication** can have biggest impact

Privacy

Age Proofs



- Pseudonym derived from certified (but secret) UID: $nym = F(uid)$
  Establishes verifiable and unforgeable pseudonymous identity
  Pseudonymous authentication is unlinkable (with ZKPs)
- Pseudonyms can additionally be scope-exclusive:
  Unique nym per scope/service, but unlinkable across
  Ensuring that users can only have single account

EUDI | Potential for Pseudonymous Authentication

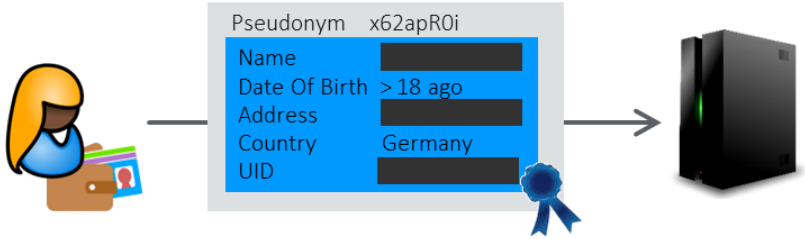
Security

Know-Your-Customer
open bank account

This is where **pseudonymous authentication** can have biggest impact

Privacy

Age Proofs

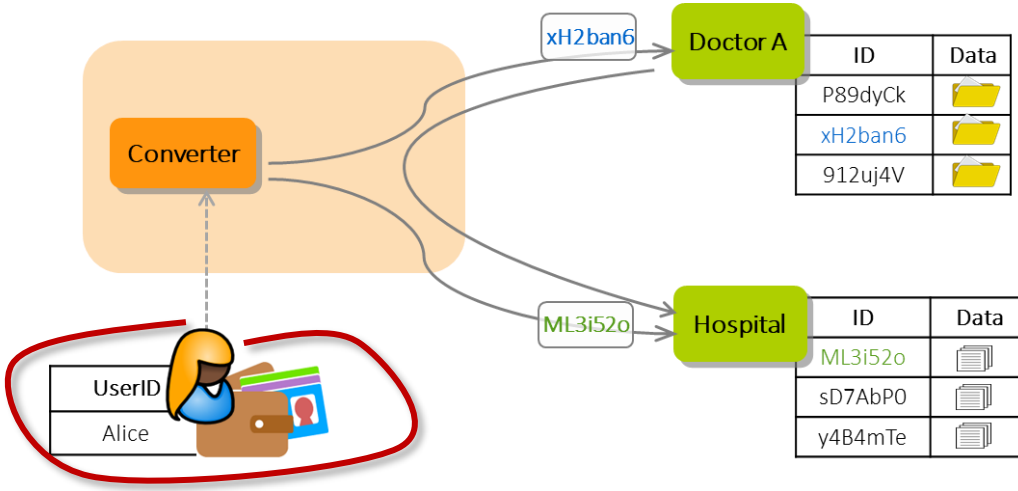


- Pseudonym derived from certified (but secret) UID: $nym = F(uid)$
  Establishes verifiable and unforgeable pseudonymous identity
  Pseudonymous authentication is unlinkable (with ZKPs)
- Pseudonyms can additionally be scope-exclusive:
  Unique nym per scope/service, but unlinkable across
  Ensuring that users can only have single account
- Many properties associated with unique and full identification can be bootstrapped in pseudonymous manner

Cryptographic Pseudonyms | Summary & State of the Art

eID with proper pseudonyms would allow to bootstrap convertible pseudonym system

- Collect already pseudonymized data
- Unlinkable pseudonyms as default
- Controlled linkability only when needed

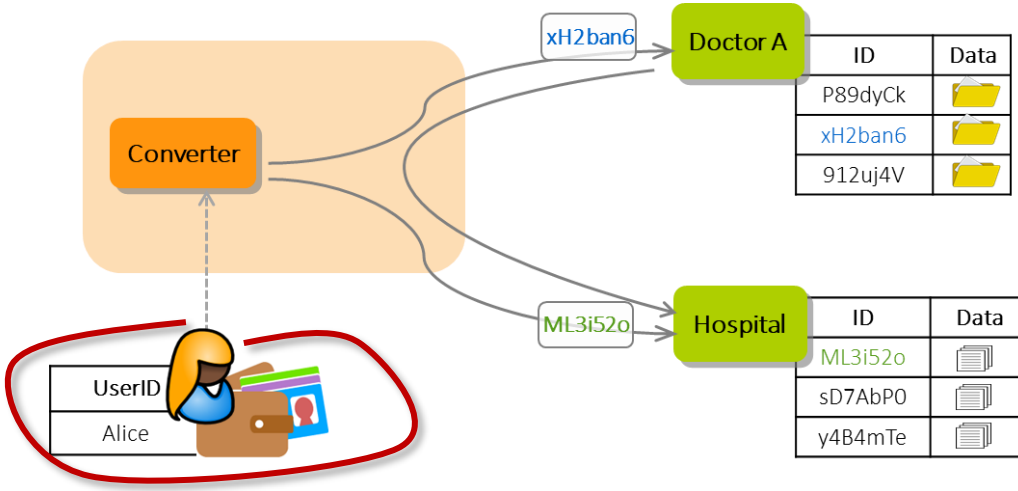


- Modern pseudonym systems enable *blind* computation & *blind* linkage of pseudonyms

Cryptographic Pseudonyms | Summary & State of the Art

eID with proper pseudonyms would allow to bootstrap convertible pseudonym system

- Collect already pseudonymized data
- Unlinkable pseudonyms as default
- Controlled linkability only when needed



- Modern pseudonym systems enable *blind* computation & *blind* linkage of pseudonyms

Note:

- Cryptographic pseudonyms focus on identification/linkage, not the associated data
  → Extra measures to ensure that associated data is not identifying (in particular when linked)

References

[CL15] *Privacy for Distributed Databases via (Un)linkable Pseudonyms*. Camenisch, Lehmann. ACM CCS 2015.

[VJM+16] *Polymorphic encryption and pseudonymisation for personalised healthcare (PEP).* Verheul, Jacobs, Meijer, Hildebrandt, Ruiter. Technical Report 2016.

[CL17] *Privacy-Preserving User-Auditable Pseudonym Systems.* Camenisch, Lehmann. IEEE EuroS&P 2017

[L19] *ScrambleDB: Oblivious (Chameleon) Pseudonymization-as-a-Service.* Lehmann. PETS 2019

[GL19] *Group Signatures with Selective Linkability.* Garms, Lehmann. PKC 2019

[MLS25] *A Brief Note on Cryptographic Pseudonyms for Anonymous Credentials.* Mayrhofer, Lehmann, shelat. https://arxiv.org/pdf/2510.05419 2025

[LMS26] *Multi-Party Private Join*. Lehmann, Mouchet, Sidorenko. PETS 2026