Privacy in Practice with Smart Pseudonymization
Lessons from the Belgian Public Sector

Kristof Verslype
Cryptographer (PhD.) @ Smals Research

15 October 2024

SAI

Smals
ICT for society

Smals
ICT for society

SUPPORT FOR E-GOVERNMENT

Knowhow

Development

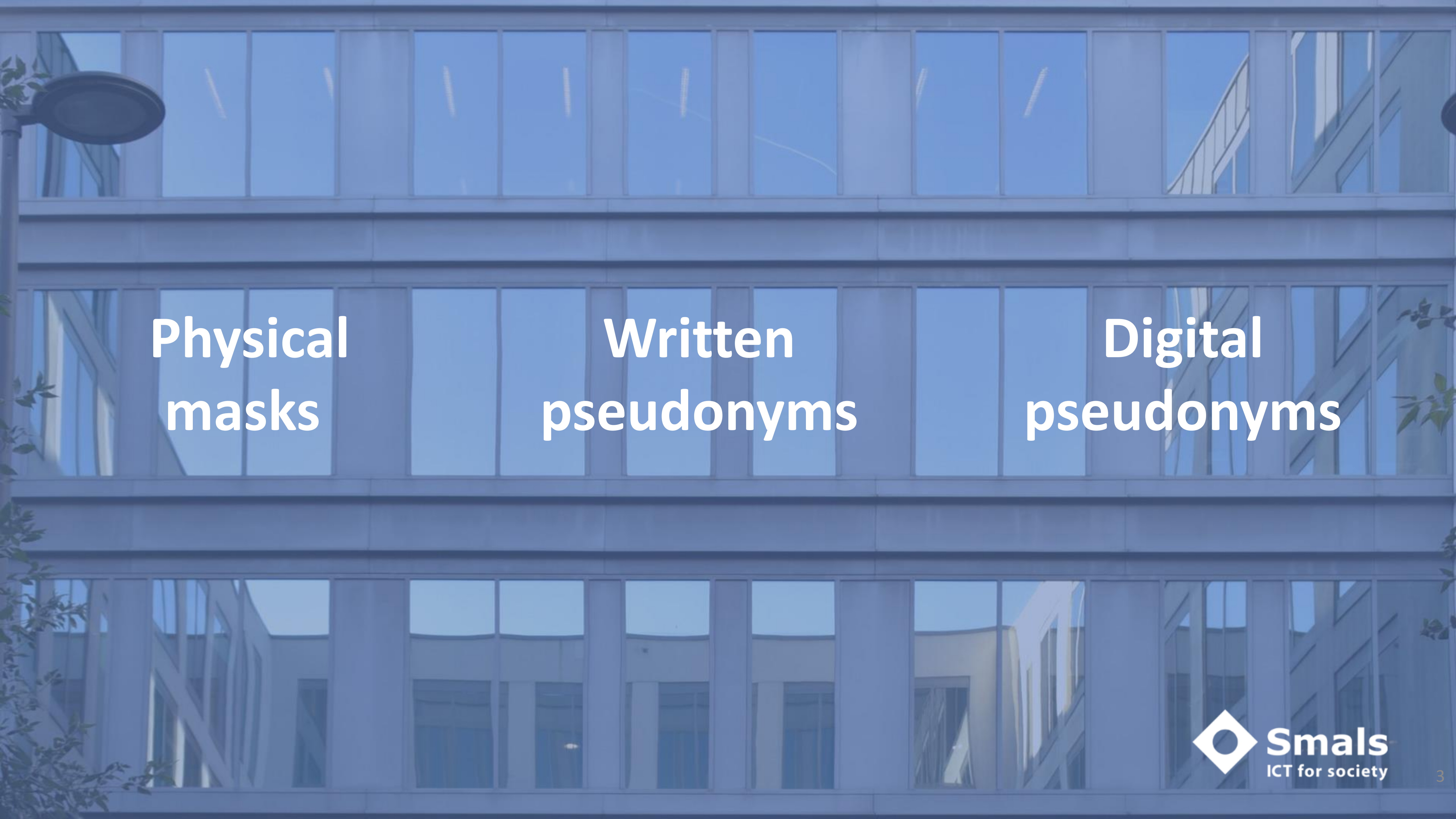
Staffing

Infrastructure

g-cloud

top EMPLOYER
BELGIË
BELGIQUE
CERTIFIED EXCELLENCE IN EMPLOYEE CONDITIONS

Smals
ICT for society
www.smals.be



Physical masks

Written pseudonyms

Digital pseudonyms

Innovation @ Smals Research

Smart Pseudonymisation

Conversion from citizen identifiers to pseudonyms

Format-Preserving Pseudonymisation

Retroactive protection of personal data in TEST & ACC of legacy applications

eHealth Blind Pseudonymisation

Proactive protection of personal data in applications
Privacy by Design

Oblivious Join

Non-trivial join & pseudonymise projects for research purposes
Distributed & no integration

ICT for society

Format-Preserving Pseudonymisation

- Problem statement
- Concept
- Experimental service
- Conclusion

Format-Preserving Pseudonymisation

- **Problem statement**
- Concept
- Experimental service
- Conclusion

Smals
ICT for society

Widespread use of personal data in non-prod environments

“60% of organisations use raw production data in test environments”

World Quality Report, 2020

Security

Data breaches from non-prod environments

2016

UBER

Hacker exploited Uber’s software development environments to break into the rideshare giant’s cloud storage

2021

T Mobile

Hacker leveraged an unprotected router to gain access to T-Mobile’s production, staging, and development servers, which compromised over 48 million social security numbers and other details.

2022

LastPass•••|

The hacker targeted the home computer of a LastPass senior DevOps engineer

No negligible risk!

Compliance with GDPR

Personal data in TEST/ACC

- ***Legal basis***
  - Informed and actively given consent?
  - Legitimate interest (gerechtvaardigd belang) of organisation?
  - Special categories of personal data
    Minors, medical data, sexual orientation, criminal data, …
  - Other legal basis?
- ***Appropriate measures***
  - Security TEST < PROD/ACC

Pseudonimisation

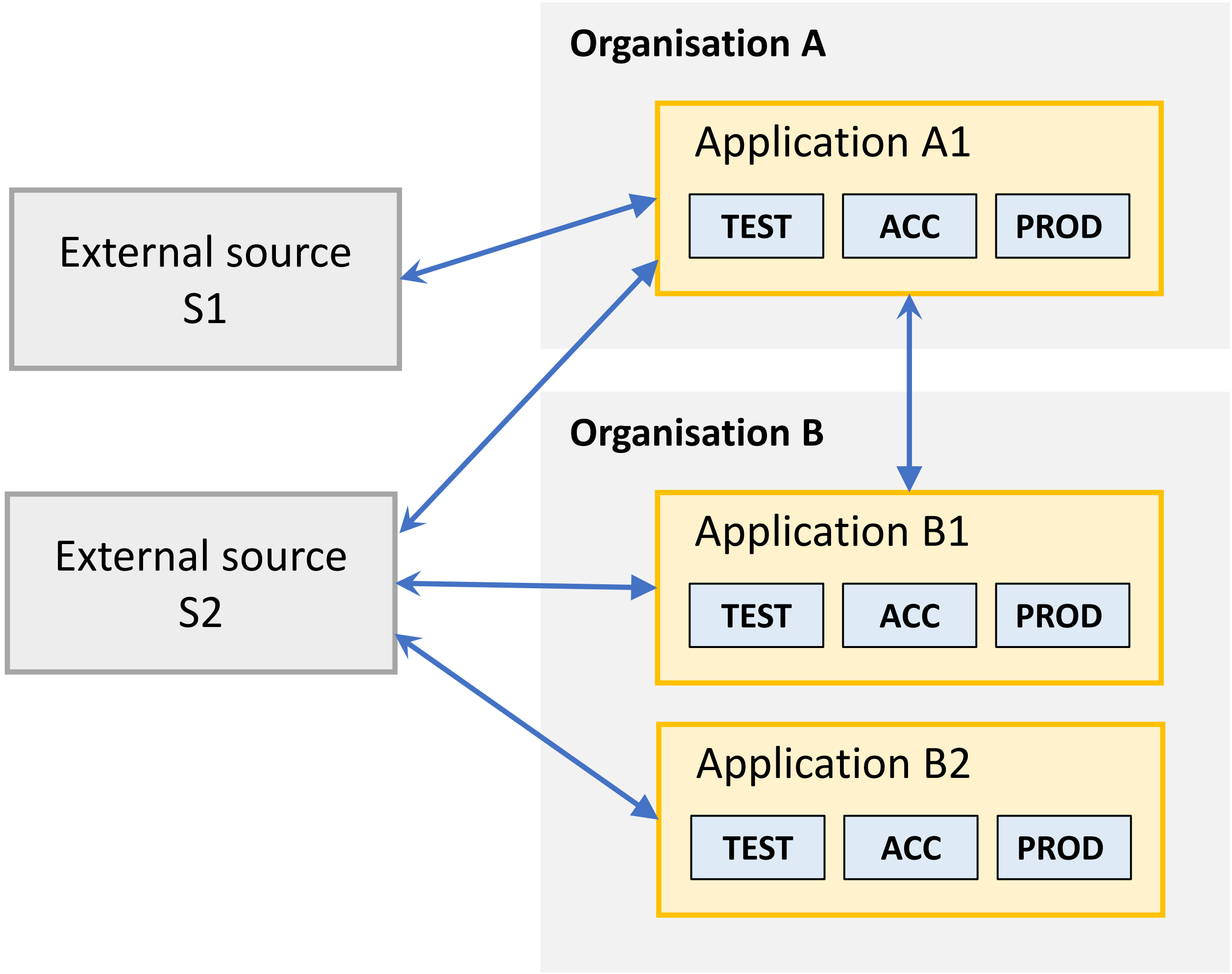
- Encouraged by GDPR to protect personal data
- Some rules by GDPR more **relaxed**
- Could help become more compliant

GDPR, Art 32.
[…] *the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:*

*a) the **pseudonymisation** and encryption of personal data;*

b) […]

Reality in public sector



Question customer

How to improve privacy in TEST & ACC?

Completely fictional data **not** an option

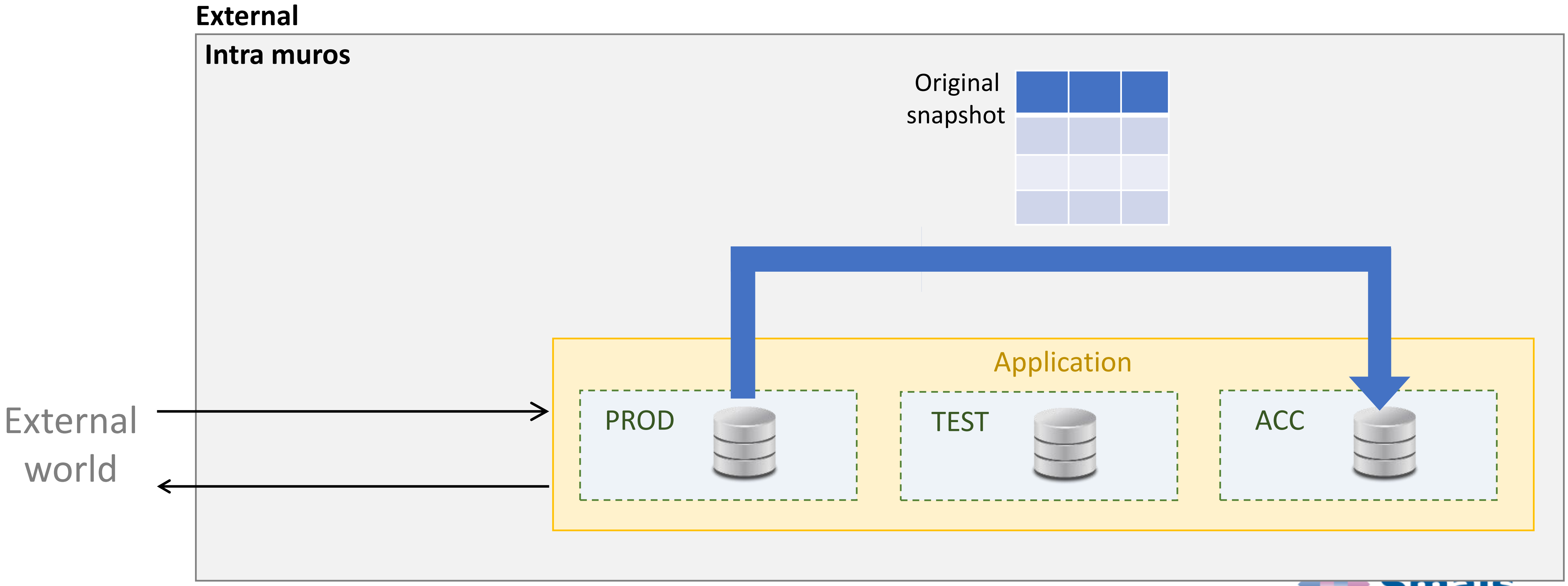
Because

- Exchange personal data
- We would miss edge cases
- Complex data models
- Labor intensive = expensive

Format-Preserving Pseudonymisation

- Problem statement
- **Concept**
- Experimental service
- Conclusion

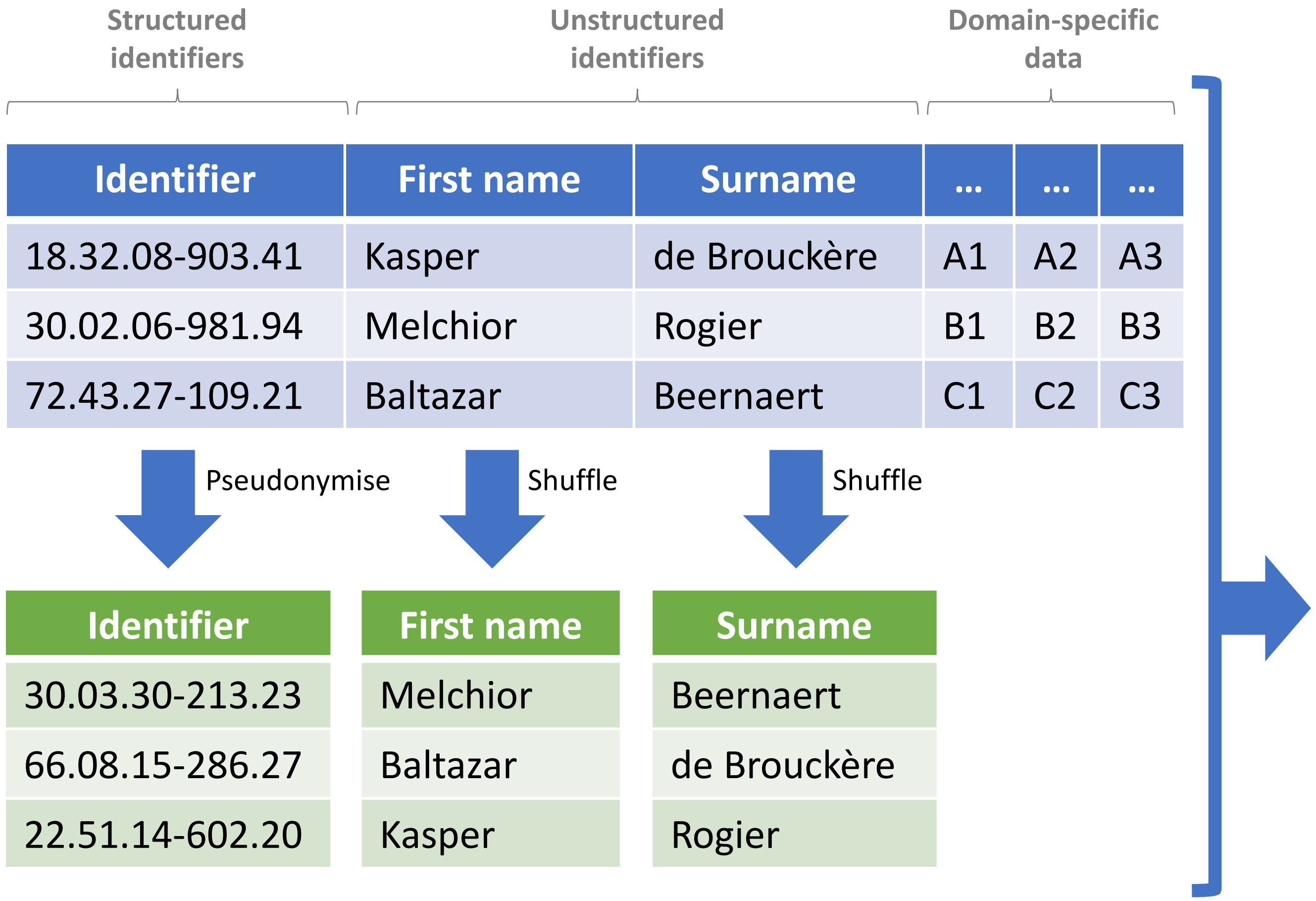
Current practice



Approach by member

Transforming batch of records with personal data copied to TEST or ACC

1. **Pseudonymise**
   Replace structured identifier by format-preserving pseudonym
   - **Bidirectional**
   - **By Smals Research**
2. **Shuffle**
   Column-wise permutation
   - **Unidirectional**
   - **By Customer**



Structured identifiers | Unstructured identifiers | Domain-specific data

Identifier	First name	Surname	…	…	…
18.32.08-903.41	Kasper	de Brouckère	A1	A2	A3
30.02.06-981.94	Melchior	Rogier	B1	B2	B3
72.43.27-109.21	Baltazar	Beernaert	C1	C2	C3

Identifier
30.03.30-213.23
66.08.15-286.27
22.51.14-602.20

First name
Melchior
Baltazar
Kasper

Surname
Beernaert
de Brouckère
Rogier

Transformed snapshot

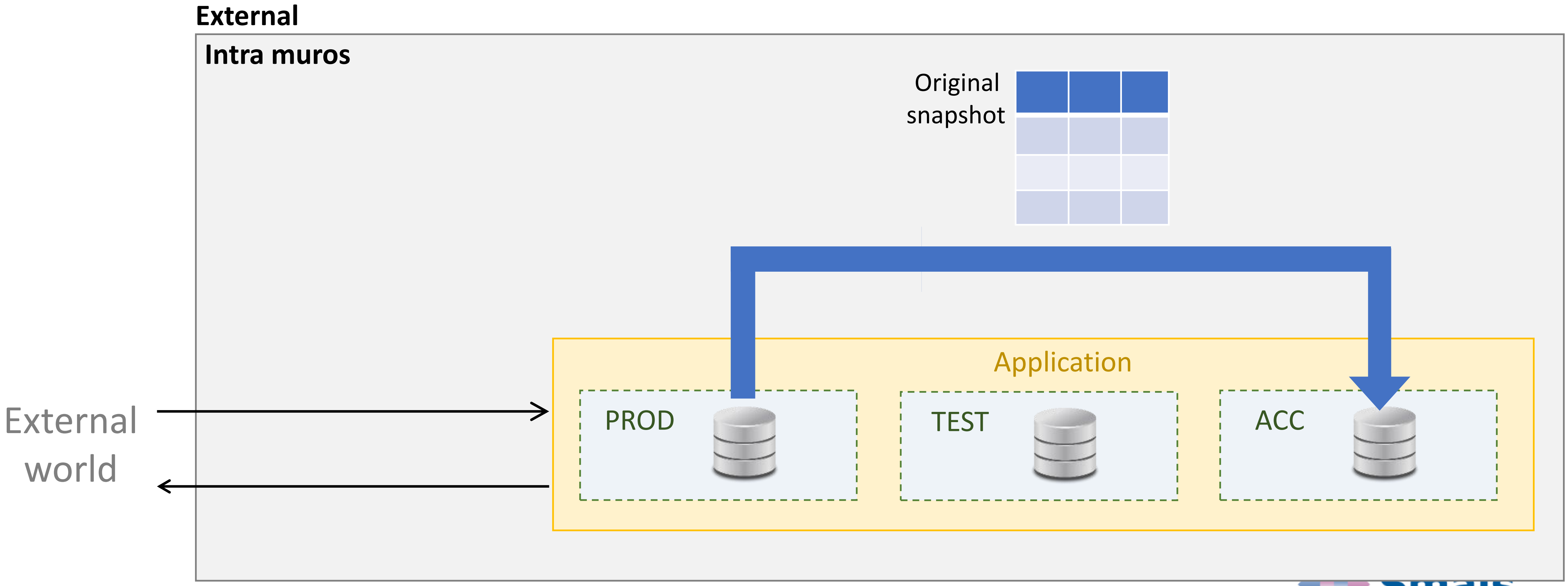
Identifier	First name	Surname	…	…	…
30.03.30-213.23	Melchior	Beernaert	A1	A2	A3
66.08.15-286.27	Baltazar	de Brouckère	B1	B2	B3
22.51.14-602.20	Kasper	Rogier	C1	C2	C3

Records useful for TEST & ACC, while hard to identify!

PoC in collaboration with customer

🟩 PoC Smals Research

⬚ PoC member



External

Intra muros

Original snapshot

Application

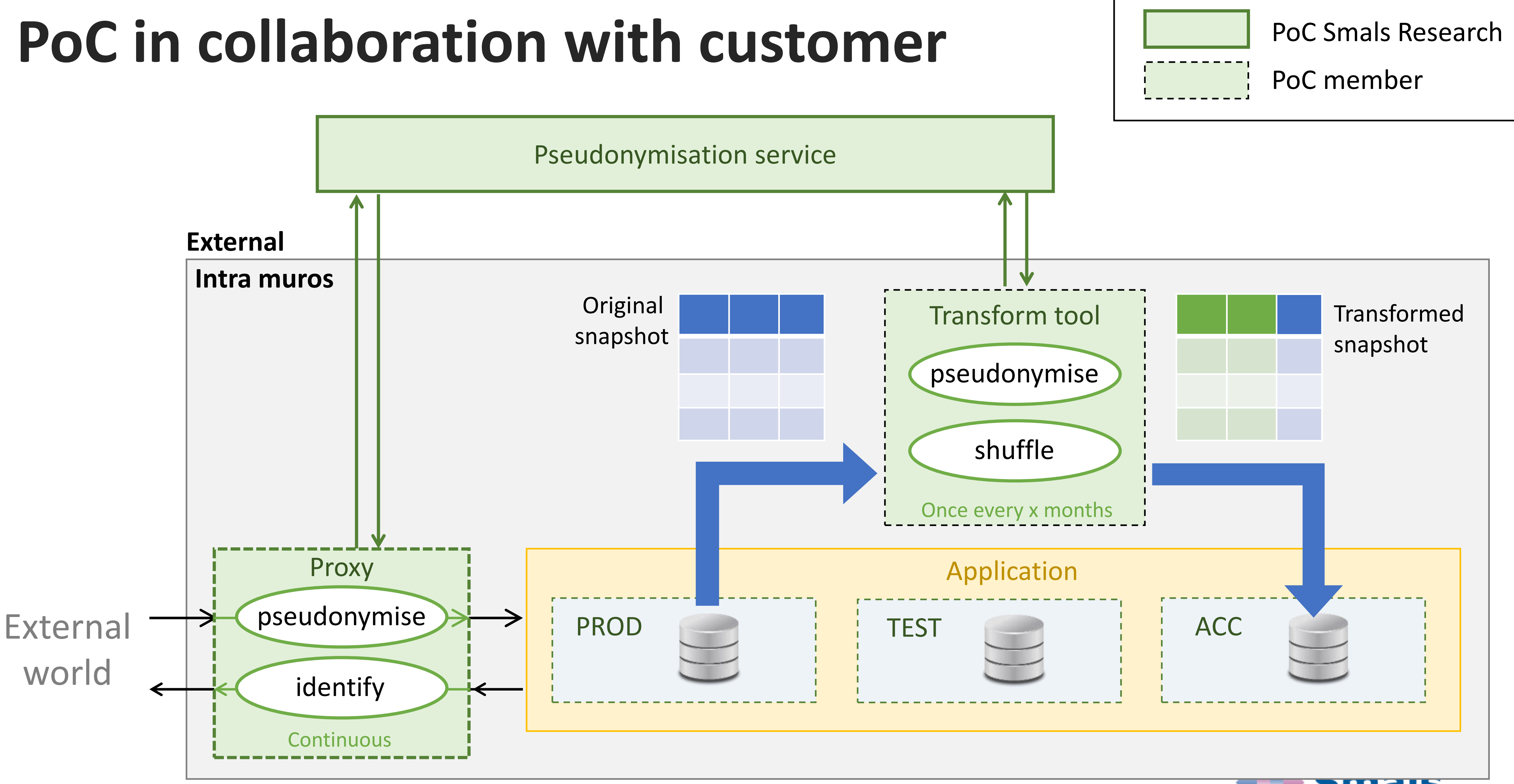
PROD

TEST

ACC

External world

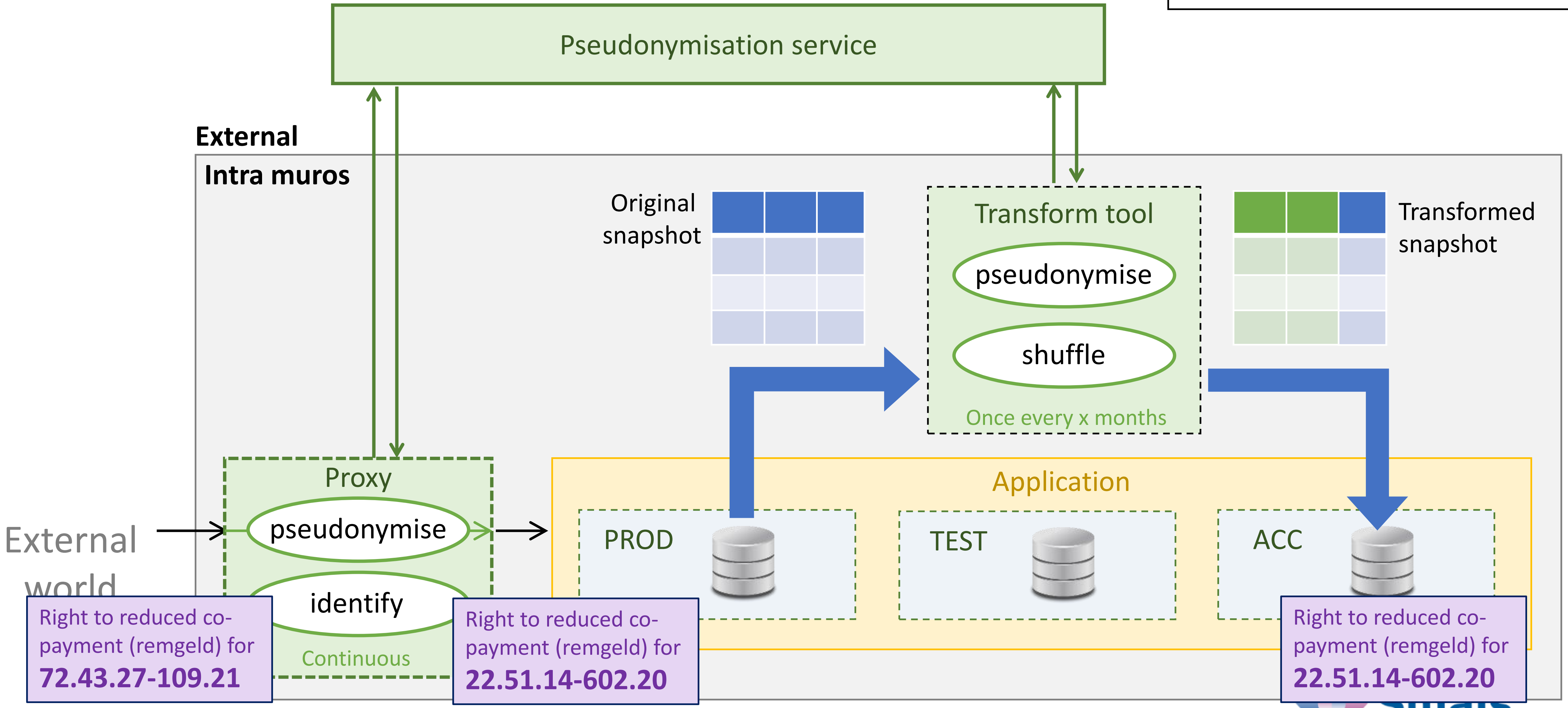
PoC in collaboration with customer



PoC in collaboration with customer

🟩 PoC Smals Research

▭ PoC member



Pseudonymisation service

External

Intra muros

Original snapshot

Transform tool

pseudonymise

shuffle

Once every x months

Transformed snapshot

Proxy

pseudonymise

identify

Continuous

Application

PROD

TEST

ACC

External world

Right to reduced co-payment (remgeld) for **72.43.27-109.21**

Right to reduced co-payment (remgeld) for **22.51.14-602.20**

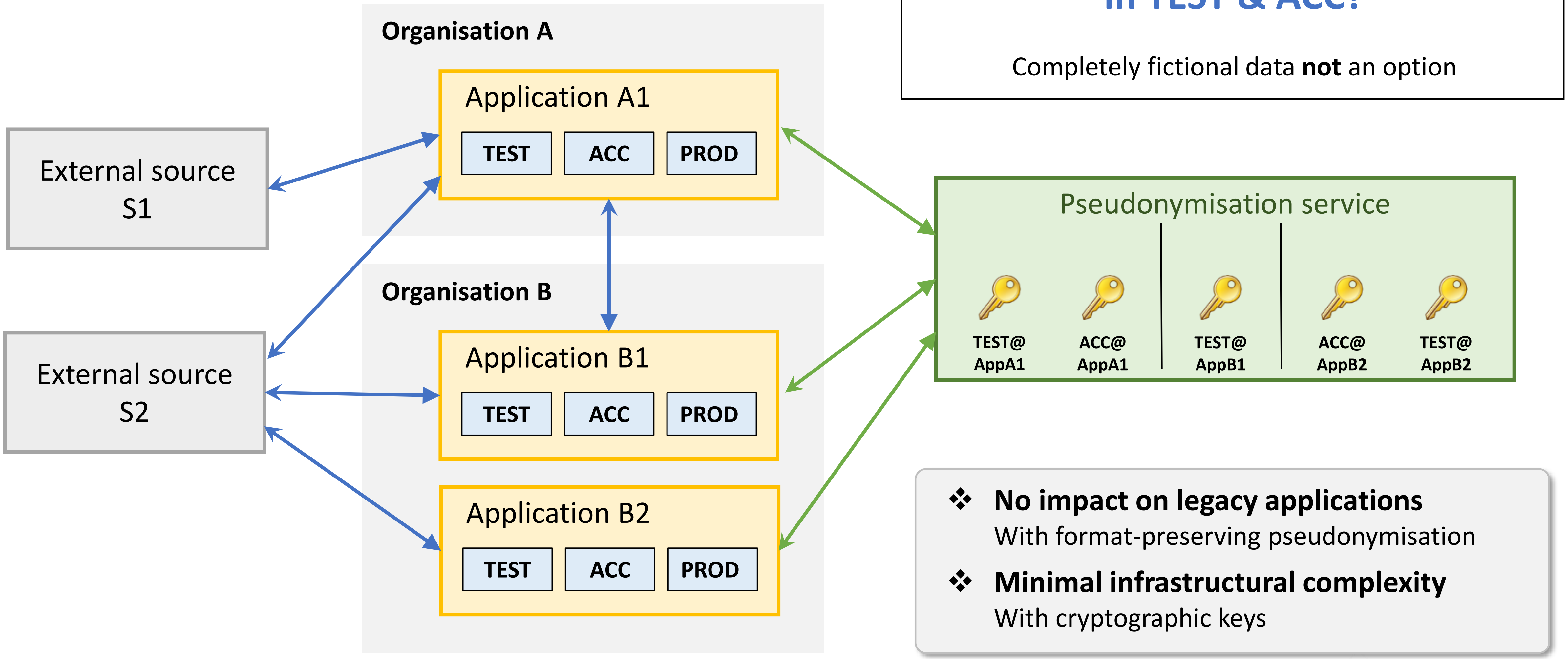
Right to reduced co-payment (remgeld) for **22.51.14-602.20**

Reality in public sector

Question customer

How to improve privacy in TEST & ACC?

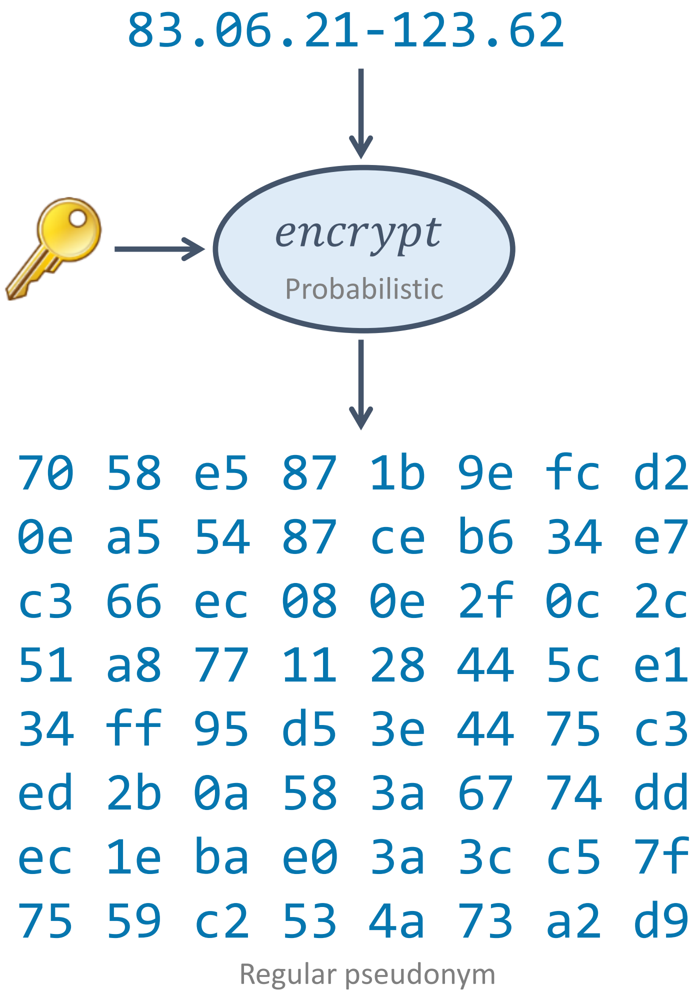
Completely fictional data **not** an option



- ❖ **No impact on legacy applications**
  With format-preserving pseudonymisation
- ❖ **Minimal infrastructural complexity**
  With cryptographic keys

Encryption

Traditional encryption

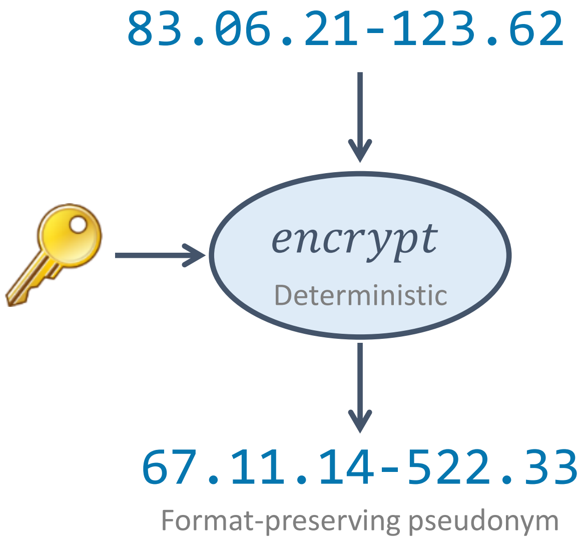


83.06.21-123.62

```
70 58 e5 87 1b 9e fc d2
0e a5 54 87 ce b6 34 e7
c3 66 ec 08 0e 2f 0c 2c
51 a8 77 11 28 44 5c e1
34 ff 95 d5 3e 44 75 c3
ed 2b 0a 58 3a 67 74 dd
ec 1e ba e0 3a 3c c5 7f
75 59 c2 53 4a 73 a2 d9
```

Regular pseudonym

Format-preserving encryption



83.06.21-123.62

67.11.14-522.33

Format-preserving pseudonym

- ❖ Conversions happen on-the-fly
- ❖ Structure preserved, including valid checksum
- ❖ Described in NIST SP 800-38G Revision. 1 (2019)

Format-Preserving Pseudonymisation

- Problem statement
- Concept
- **Experimental service**
- Conclusion

Smals
ICT for society

Experimental REST service

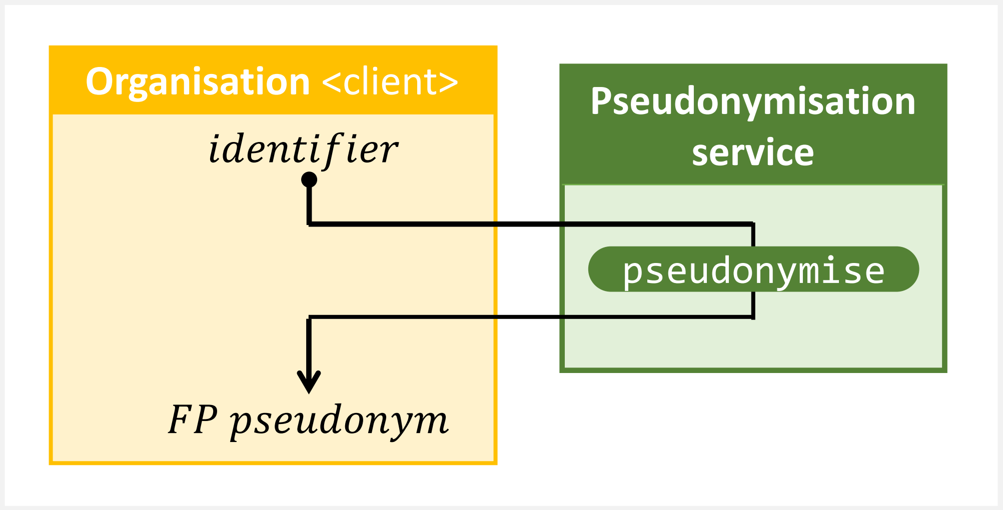
Built by Smals Research

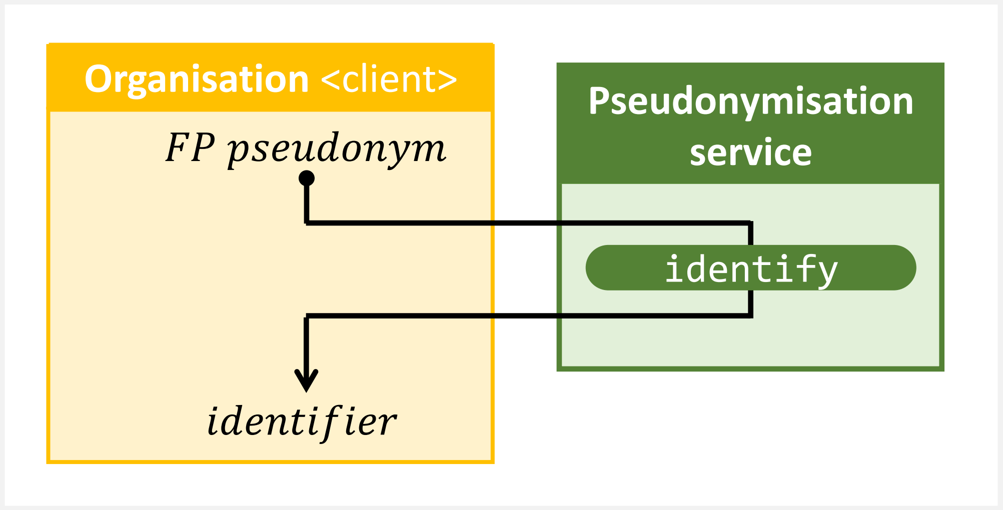
Rest API

- ✓ *Pseudonymise & Identify*
- ✓ *GET and POST*
- ✓ *Also batch (POST only)*

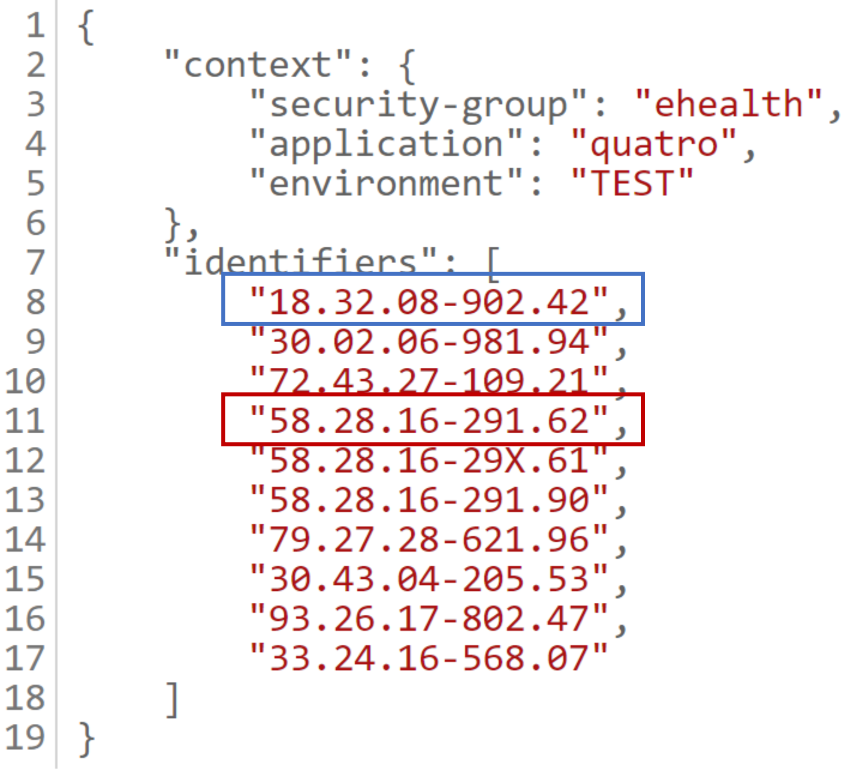
Identifiers

- ✓ Support for Belgian social security numbers
- ✓ *Extensible*





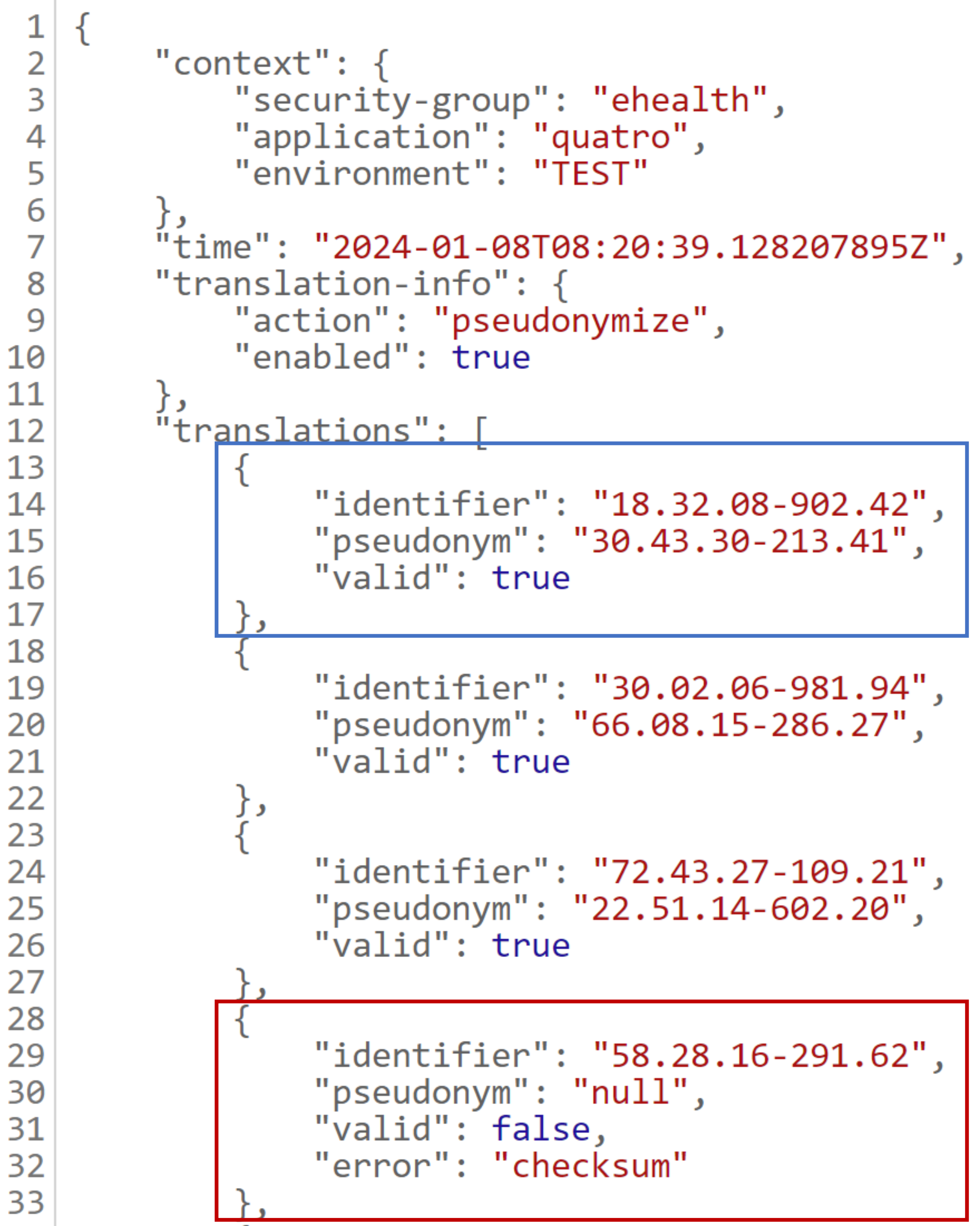
POST Request



```
{
    "context": {
        "security-group": "ehealth",
        "application": "quatro",
        "environment": "TEST"
    },
    "identifiers": [
        "18.32.08-902.42",
        "30.02.06-981.94",
        "72.43.27-109.21",
        "58.28.16-291.62",
        "58.28.16-29X.61",
        "58.28.16-291.90",
        "79.27.28-621.96",
        "30.43.04-205.53",
        "93.26.17-802.47",
        "33.24.16-568.07"
    ]
}
```

- ✓ Easy to use
- ✓ Graceful error handling
- ✓ Efficient

POST Response



```
{
    "context": {
        "security-group": "ehealth",
        "application": "quatro",
        "environment": "TEST"
    },
    "time": "2024-01-08T08:20:39.128207895Z",
    "translation-info": {
        "action": "pseudonymize",
        "enabled": true
    },
    "translations": [
        {
            "identifier": "18.32.08-902.42",
            "pseudonym": "30.43.30-213.41",
            "valid": true
        },
        {
            "identifier": "30.02.06-981.94",
            "pseudonym": "66.08.15-286.27",
            "valid": true
        },
        {
            "identifier": "72.43.27-109.21",
            "pseudonym": "22.51.14-602.20",
            "valid": true
        },
        {
            "identifier": "58.28.16-291.62",
            "pseudonym": "null",
            "valid": false,
            "error": "checksum"
        },
```

Format-Preserving Pseudonymisation

- Problem statement
- Concept
- Experimental service
- **Conclusion**

Format-Preserving Pseudonymisation

Building block

- To improve privacy in TEST and ACC environments
- Partial solution

As a Service

- Simplifies logic organisation
  E.g. key management
- Stimulates reuse
- Separation of duties

Status

Trying to go into project mode

Innovation @ Smals Research

Smart Pseudonymisation

Conversion from citizen identifiers to pseudonyms

Format-Preserving Pseudonymisation

Retroactive protection of personal data in TEST & ACC of legacy applications

eHealth Blind Pseudonymisation

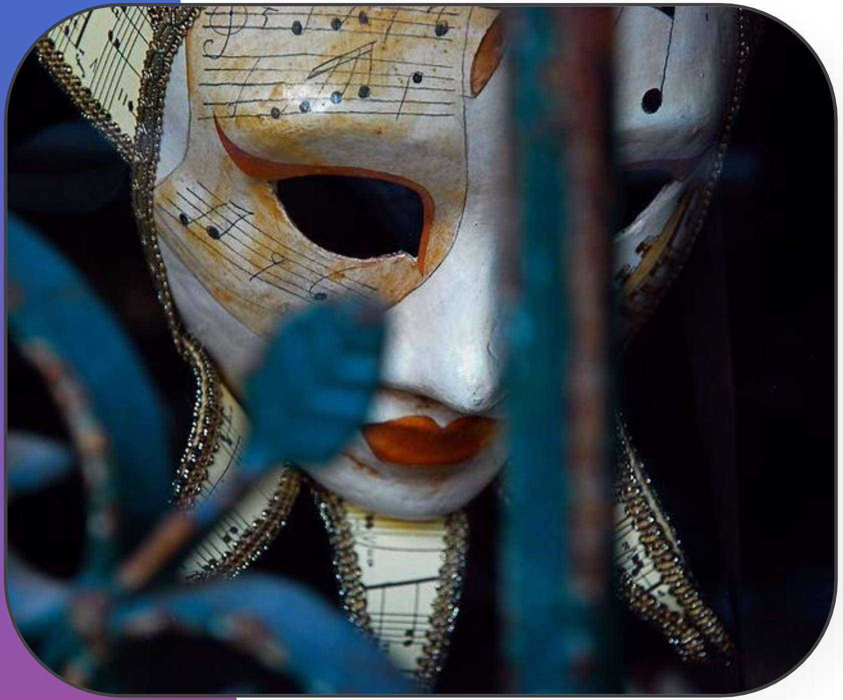
Proactive protection of personal data in applications
Privacy by Design

Oblivious Join

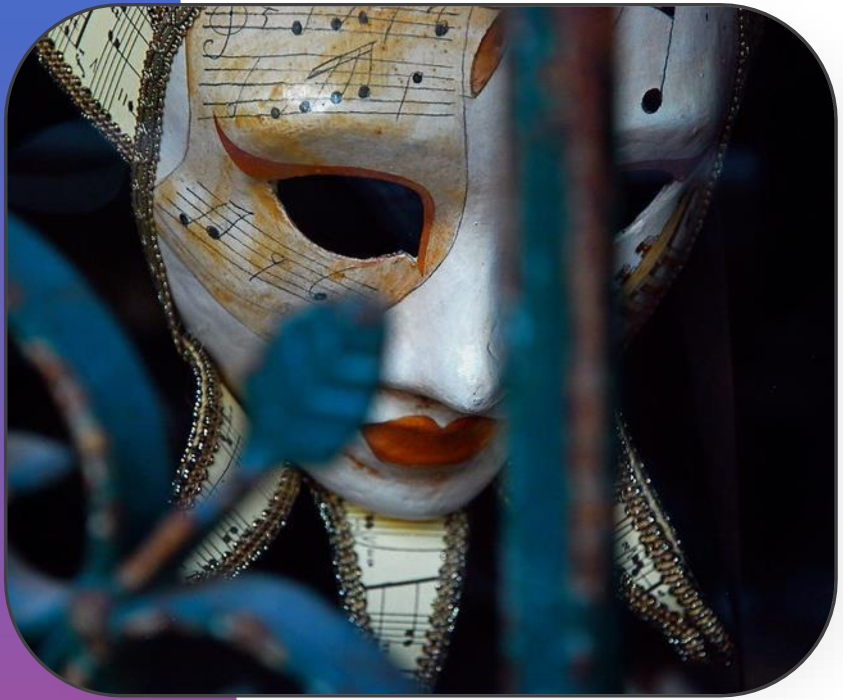
Non-trivial join & pseudonymise projects for research purposes
Distributed & no integration

ICT for society

eHealth Blind Pseudonymisation



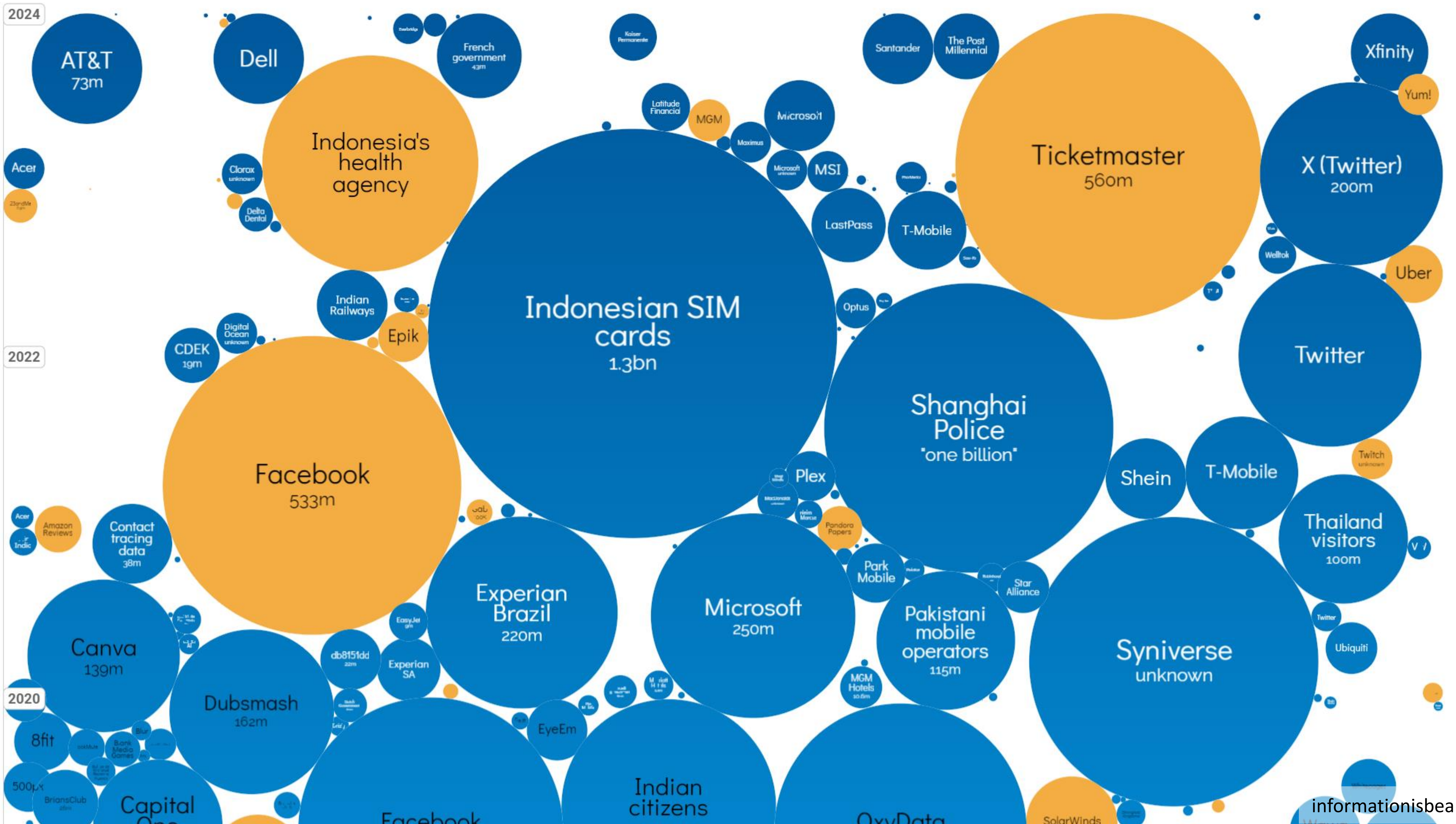
- Problem statement
- Referral prescriptions
- Join & pseudonymise data for research
- Conclusion



eHealth Blind Pseudonymisation

- **Problem statement**
- Referral prescriptions
- Join & pseudonymise data for research
- Conclusion

World's Biggest Data Breaches & Hacks



informationisbeautiful.net

Design principles

Privacy by design

Privacy should be taken into account when designing and building products and services

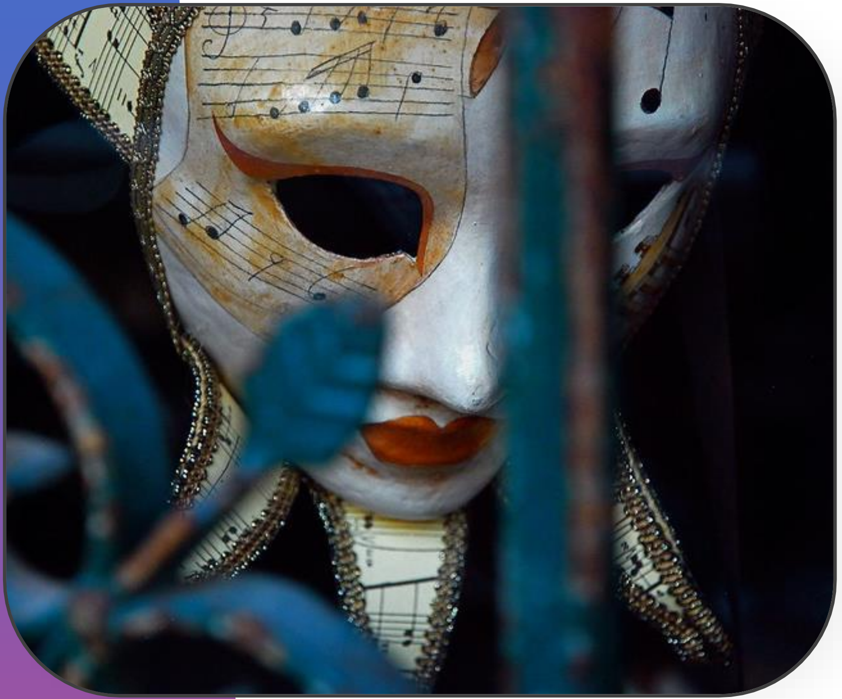
Separation of duties

Entity managing protection keys should not have access to protected data (and vice versa)

Simplicity

Complexity is the worst enemy of security

eHealth Blind Pseudonymisation



- Problem statement
- **Referral prescriptions**
- Join & pseudonymise data for research
- Conclusion

Smals
ICT for society

Use case 1 - Live

Referral prescription = Verwijsvoorschrift / Prescription de renvoi

What?

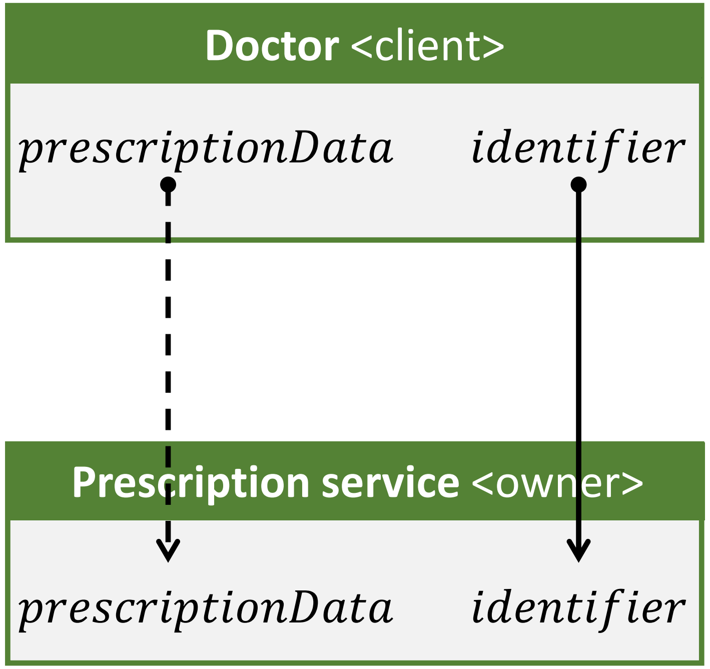
A certificate to start a certain treatment (e.g. physiotherapist, dieticians, speech therapists).

Requirements

- **Pseudonymisation**
  Prescription service should never be able to link prescription data to a citizen
- **Partial encryption**
  The prescription service should not be able to access certain fields

Scenario 1

Doctor (client) requests Prescription service (owner) to register prescription



Doctor <client>

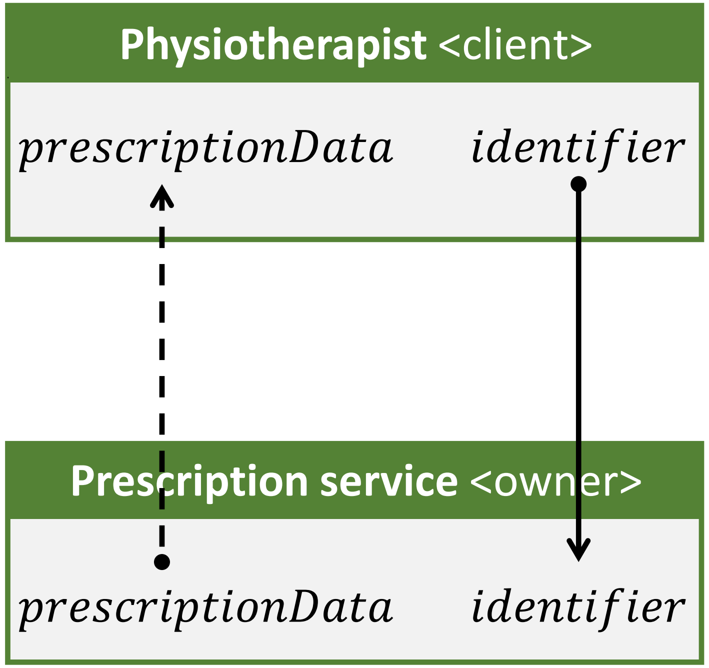
$\mathit{prescriptionData}$ $\mathit{identifier}$

Prescription service <owner>

$\mathit{prescriptionData}$ $\mathit{identifier}$

Scenario 2

Physiotherapist (client) requests access to prescription for a specific citizen from Prescription service (owner)



Physiotherapist <client>

$\mathit{prescriptionData}$ $\mathit{identifier}$

Prescription service <owner>

$\mathit{prescriptionData}$ $\mathit{identifier}$

Blind Pseudo Service

Pseudonymise

Structure blinded identifier, blinded pseudonym and final pseudonym

(AV+VXF9H5LdTe4b1 SSC7bHjp6b2enJmf plC6a3/jCR5fUHxX RSaRniYR8h7ugNqa lGvP49cZnv6lf9B7 2RUG0rA/,
eSmIl52CEtsZzSseU DY3YKLtSgqhq1wLPm 9ncHBzGiv1wMIxmc1 jSmpW36GhTt/s1P5s hZGhG8ncoWKSGkJDy fw=)

✓ **Each party only sees only what it needs to see**

- ❖ Client only sees identifiers
- ❖ Owner only sees pseudonyms
- ❖ Pseudon. service sees neither

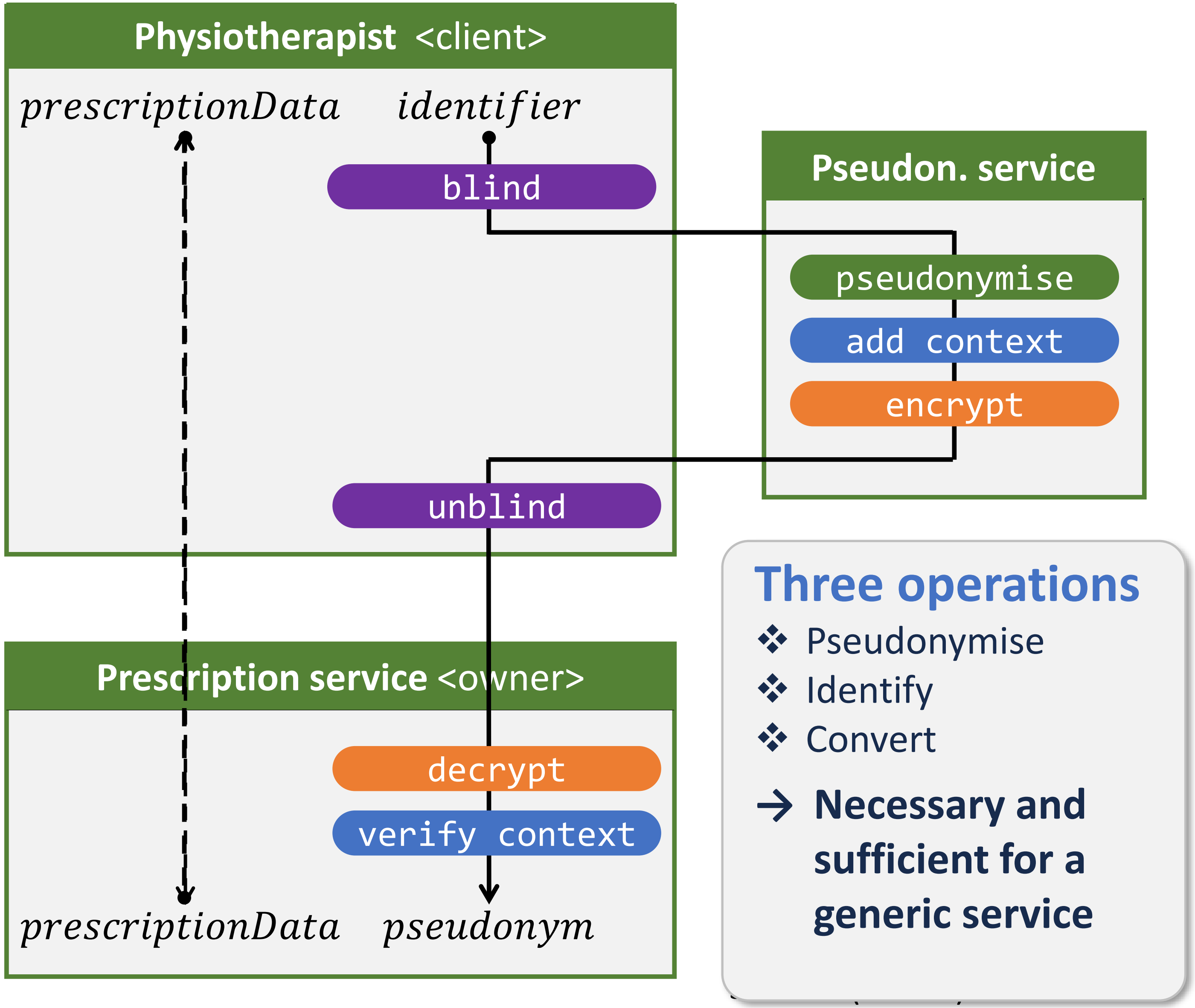
→ **Maximizes security & privacy**

✓ **Direct communication**

- ❖ Direct communication between healthcare professional and prescription service
- ❖ No in-between entity

✓ **Low-intrusive client-side**

- ❖ No extra keys required
- ❖ Relatively simple implementation



Three operations

- ❖ Pseudonymise
- ❖ Identify
- ❖ Convert

→ **Necessary and sufficient for a generic service**

Use case 1 - Live

Referral prescription = Verwijsvoorschrift / Prescription de renvoi

What?

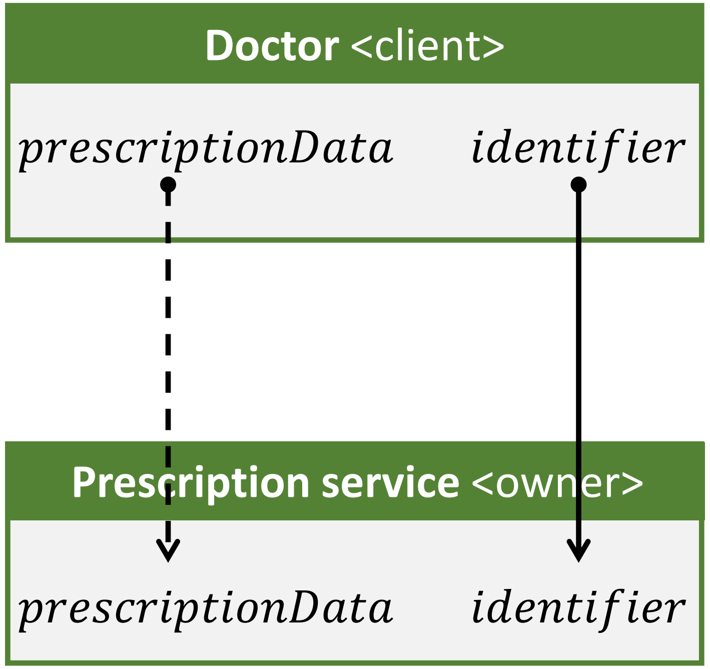
A certificate to start a certain treatment (e.g. physiotherapist, dieticians, speech therapists).

Requirements

- **Pseudonymisation**
  Prescription service should never be able to link prescription data to a citizen
- **Partial encryption**
  The prescription service should not be able to access certain fields

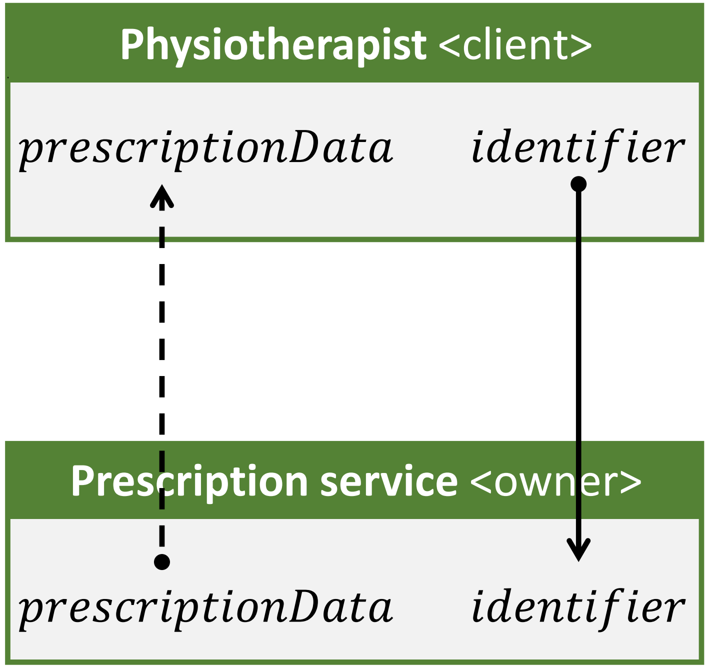
Scenario 1

Doctor (client) requests Prescription service (owner) to register prescription

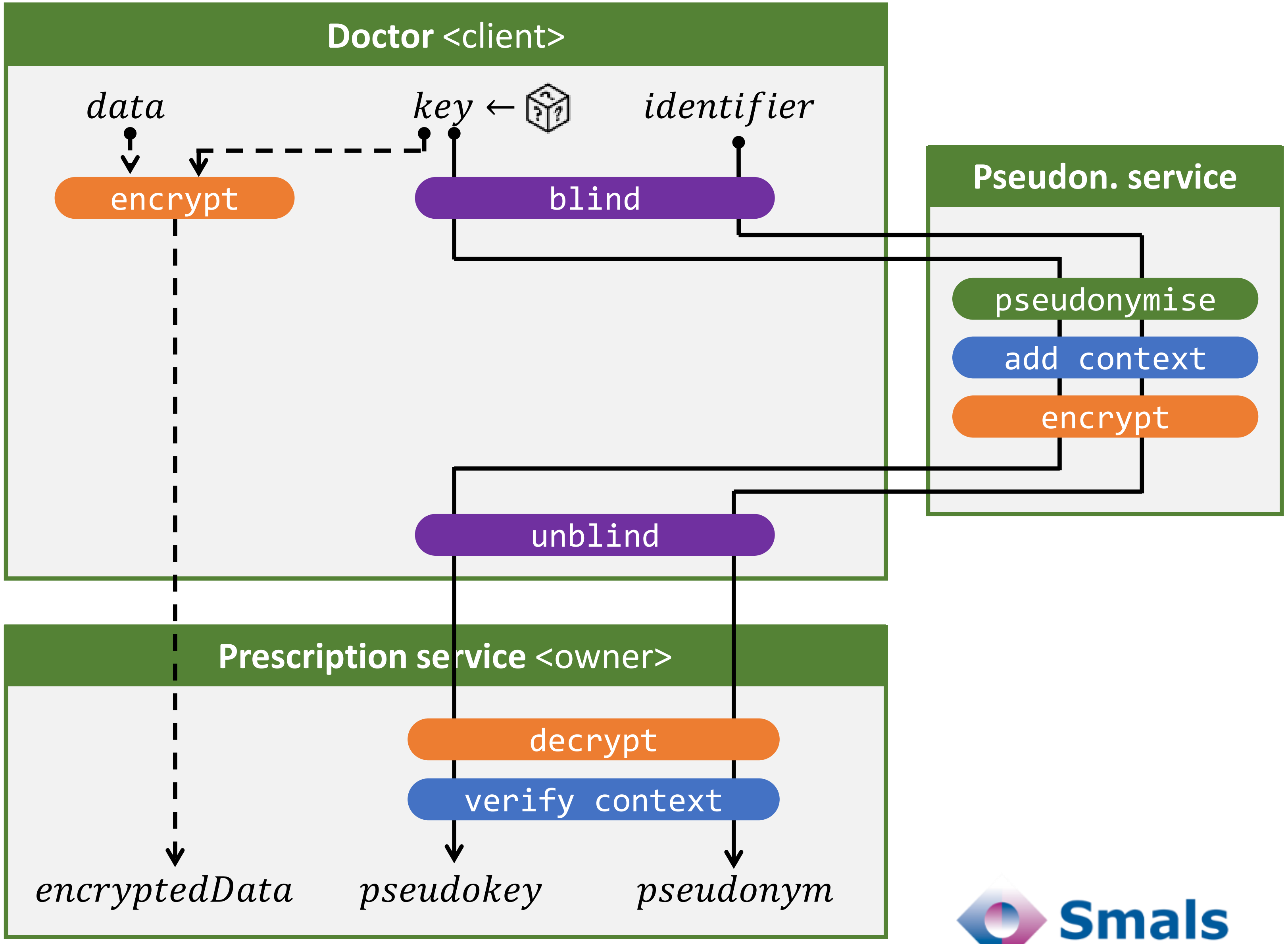


Scenario 2

Physiotherapist (client) requests access to prescription for a specific citizen from Prescription service (owner)


Physiotherapist <client>
$prescriptionData$
$identifier$
Prescription service <owner>
$prescriptionData$
$identifier$

Encrypt



Smals ICT for society

Blind Pseudonymisation Service

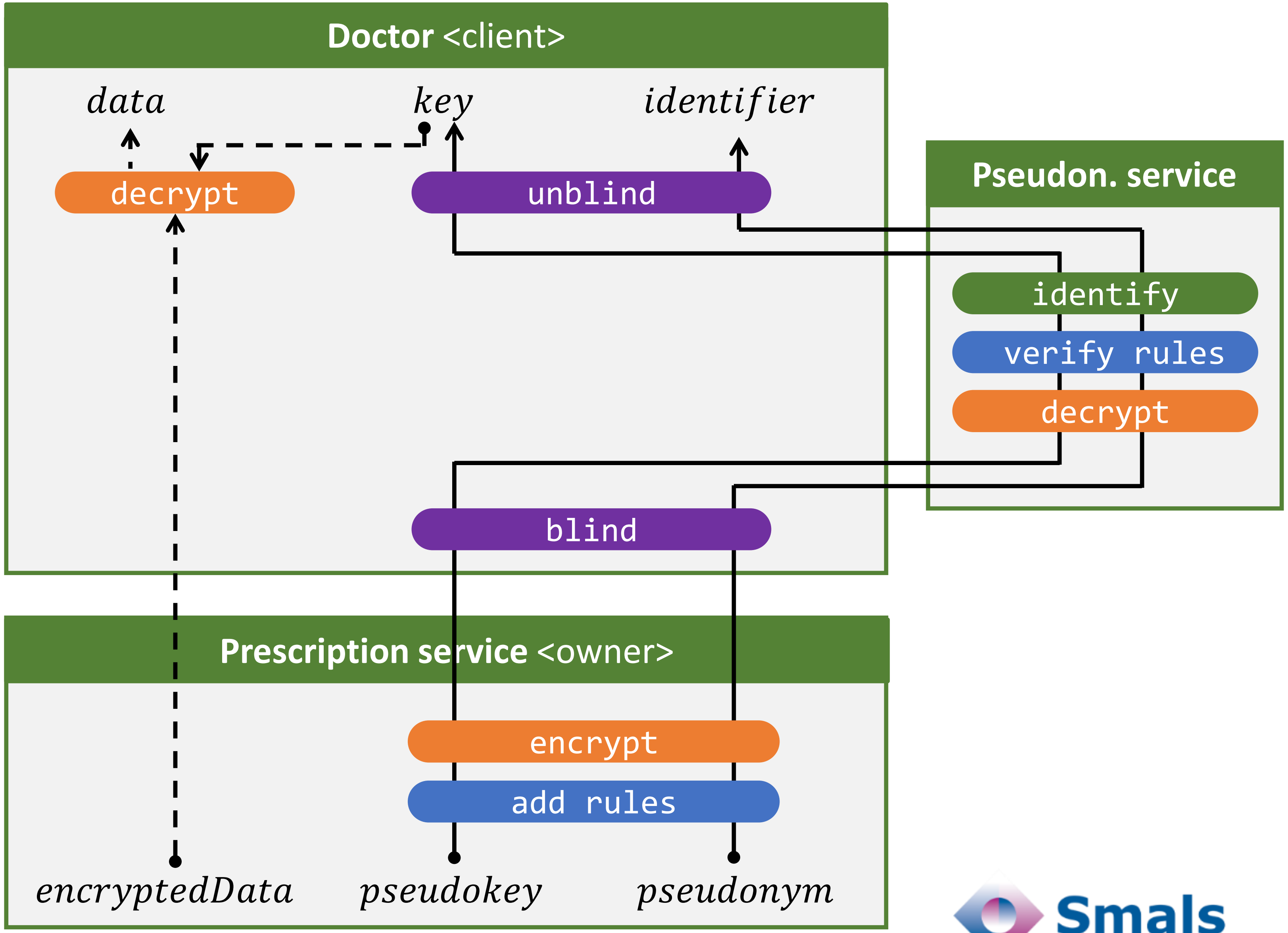
Decrypt

- ✓ Authorized healthcare professional can access data
- ✓ Prescription service cannot access data
- ✓ Pseudon. service cannot access key
- ✓ Quasi no new logic required

Crucial that pseudon. service

- ❖ is independent
- ❖ is well secured
- ❖ has proper access control

!

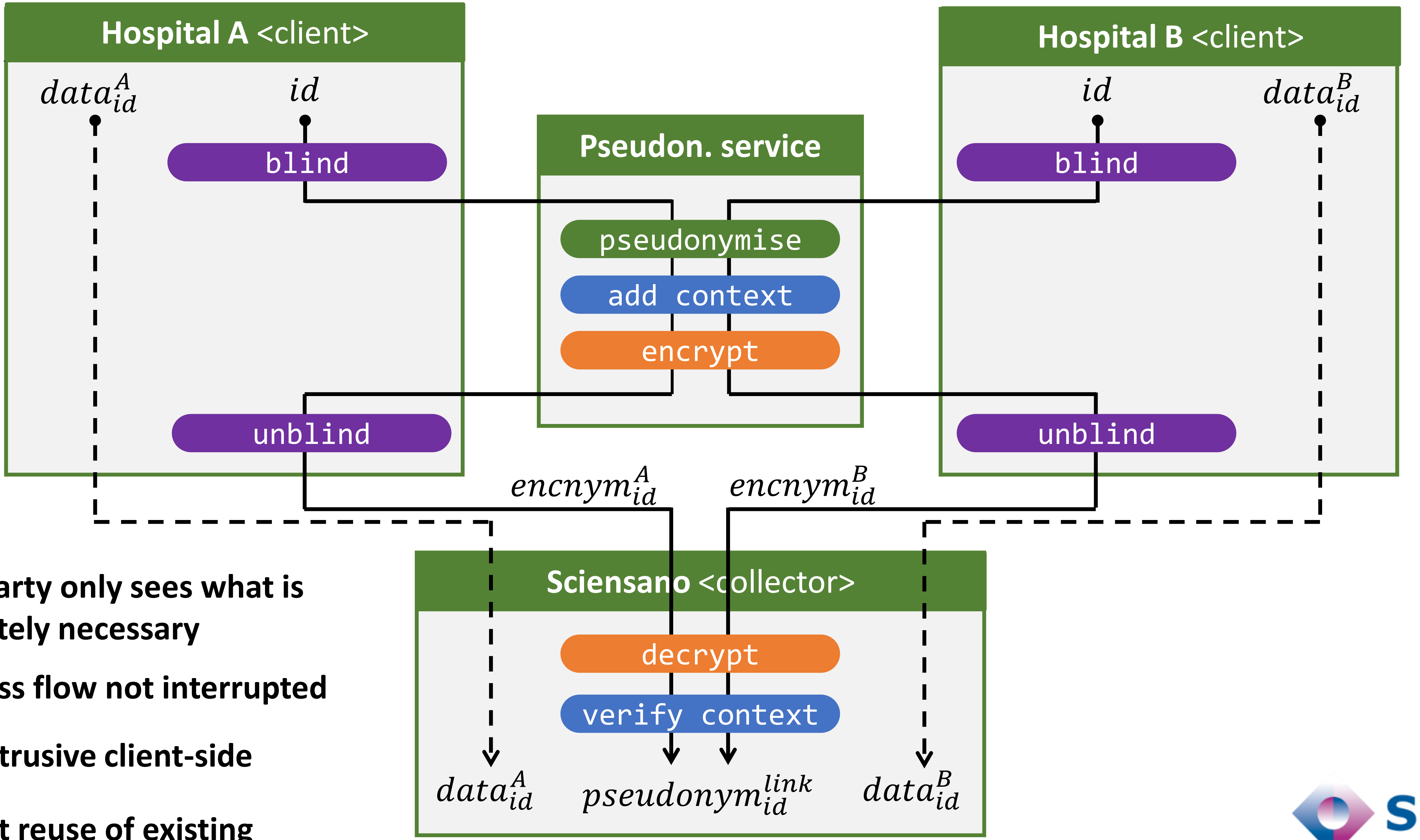


eHealth Blind Pseudonymisation

- Problem statement
- Referral prescriptions
- **Join & pseudonymise data for research**
- Conclusion

Use case 2 – Proposal

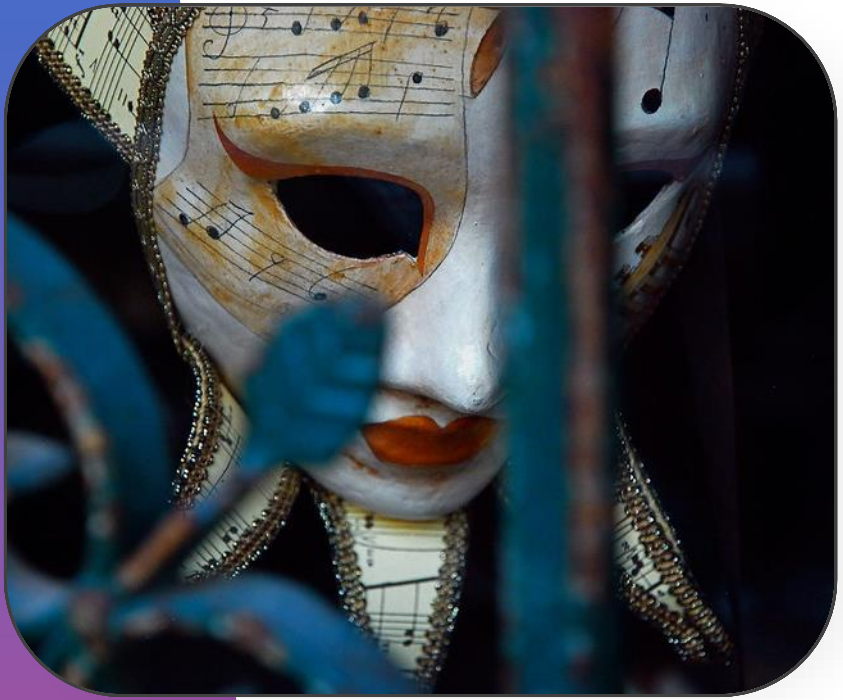
Join & pseudonymise data for research



- ✓ **Each party only sees what is absolutely necessary**
- ✓ **Business flow not interrupted**
- ✓ **Low-intrusive client-side**
- ✓ **Efficient reuse of existing infrastructure**

Smals
ICT for society

eHealth Blind Pseudonymisation



- Problem statement
- Referral prescriptions
- Join & pseudonymise data for research
- **Conclusion**

eHealth Blind Pseudonymisation

Privacy by design

- ✓ Client only sees identifiers
- ✓ Service only sees pseudons.
- ✓ Pseudon. service sees neither

Separation of duties

Pseudon. service and prescription service are separate entities

Simplicity

- ✓ Versatility
- ✓ Low complexity client slide

Live with uptake

ePrescriptions, medication regimens, medical record summaries, vaccinations, allergies and intolerances, fertility.

Smals
ICT for society

Innovation @ Smals Research

Smart Pseudonymisation

Conversion from citizen identifiers to pseudonyms

Format-Preserving Pseudonymisation

Retroactive protection of personal data in TEST & ACC of legacy applications

eHealth Blind Pseudonymisation

Proactive protection of personal data in applications
Privacy by Design

Oblivious Join

Non-trivial join & pseudonymise projects for research purposes
Distributed & no integration

ICT for society

Oblivious Join

- Problem statement
- Concept
- In practice
- Conclusion

Oblivious Join

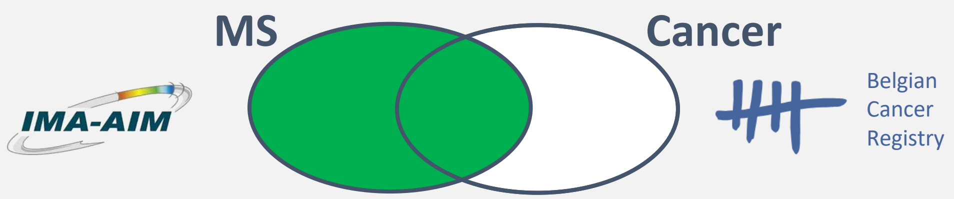
- **Problem statement**
- Concept
- In practice
- Conclusion

Concrete case

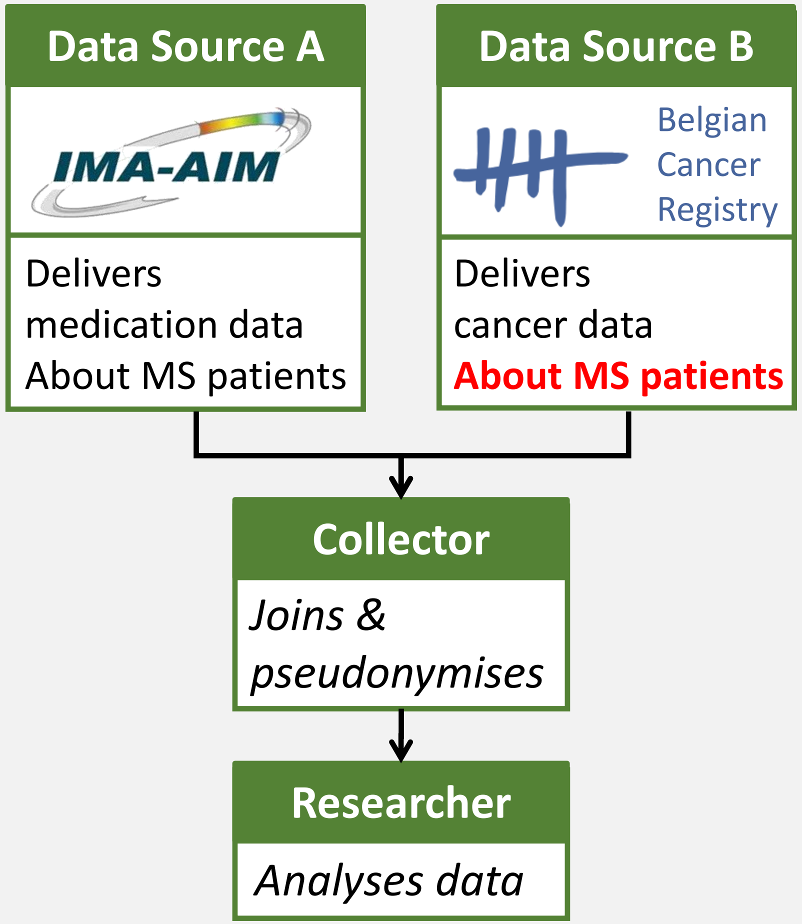
Research question

Do MS patients who take medications with the molecule teriflunomide or alemtuzumab have an increased cancer risk compared to MS patients treated with other medications?

Involved citizens



Naive flow



Data Source A

Delivers medication data About MS patients

Data Source B

Delivers cancer data **About MS patients**

Collector

Joins & pseudonymises

Researcher

Analyses data

How can BCR deliver only records about MS patients without learning who has MS?

Current practice

Observations

- ✗ Complex flow
- ✗ Bespoke
- ✗ Slow
- ✗ Expensive
- ✗ Doesn’t scale well
- ✗ Security risk (data leakage)

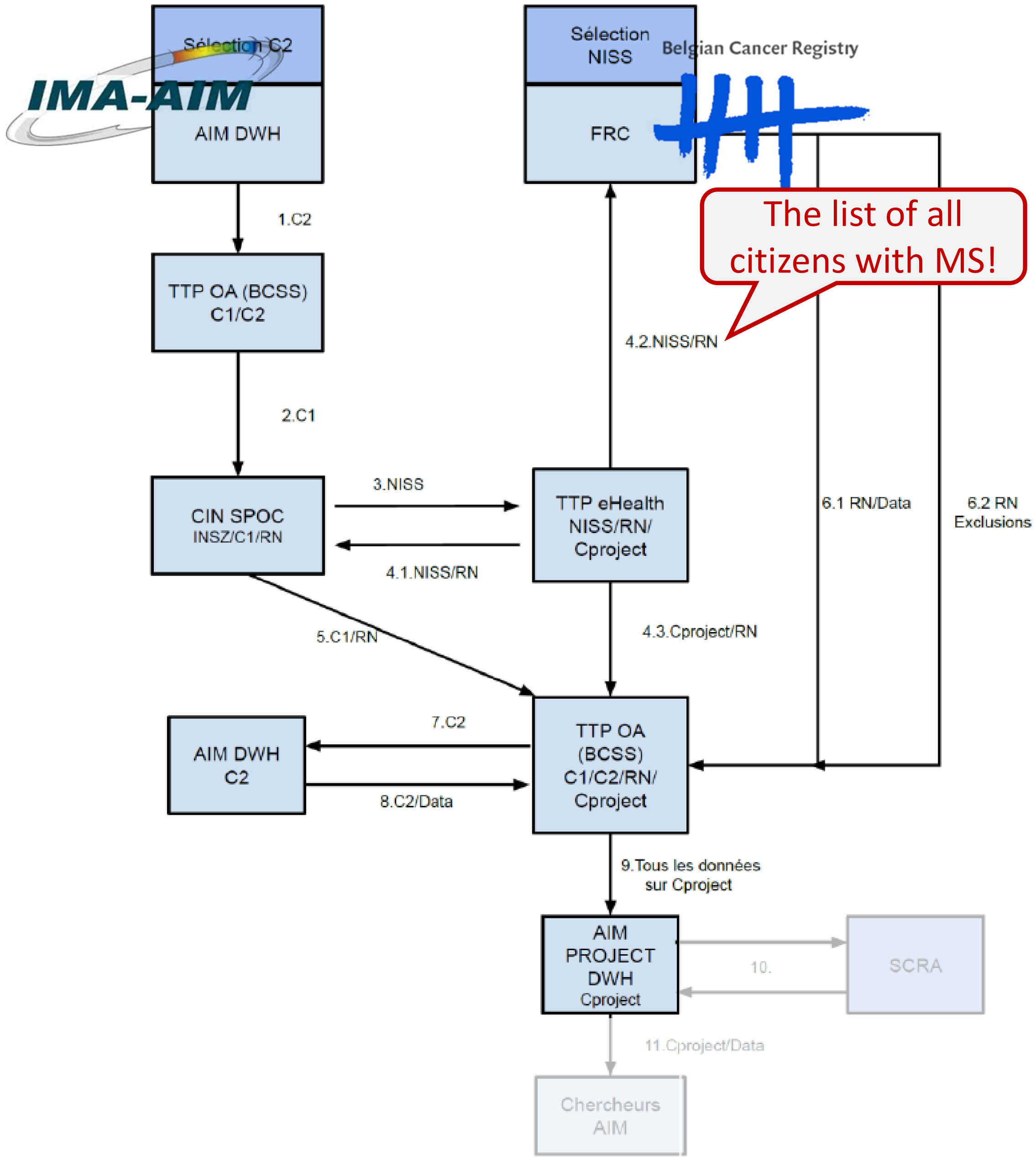
Feedback

“Lasts weeks, months, even years”

“Requires an exorbitant amount of resources”

Other countries

Heavy reliance on combination of trusted parties and strong legal regulations



The list of all citizens with MS!

Challenge

Join and pseudonymise personal data originating from different sources

Constraint

Not all data sources able to independently select relevant records

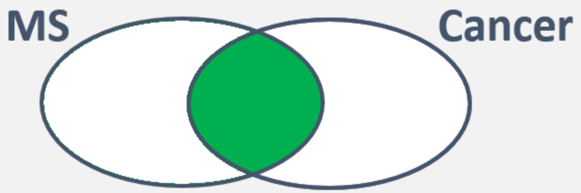
E.g., BCR unable to select records about citizens with MS

Requirements

- **Privacy-friendly**
  Involved entities learns only the necessary
- **Uniform**
  Each research question is different, with different data and different data sources
- **No data aggregation**
  Researcher access to individual records
- **Easy to use**

Focus: set intersection

Researcher wants pseudonymised data of citizens that have MS and cancer

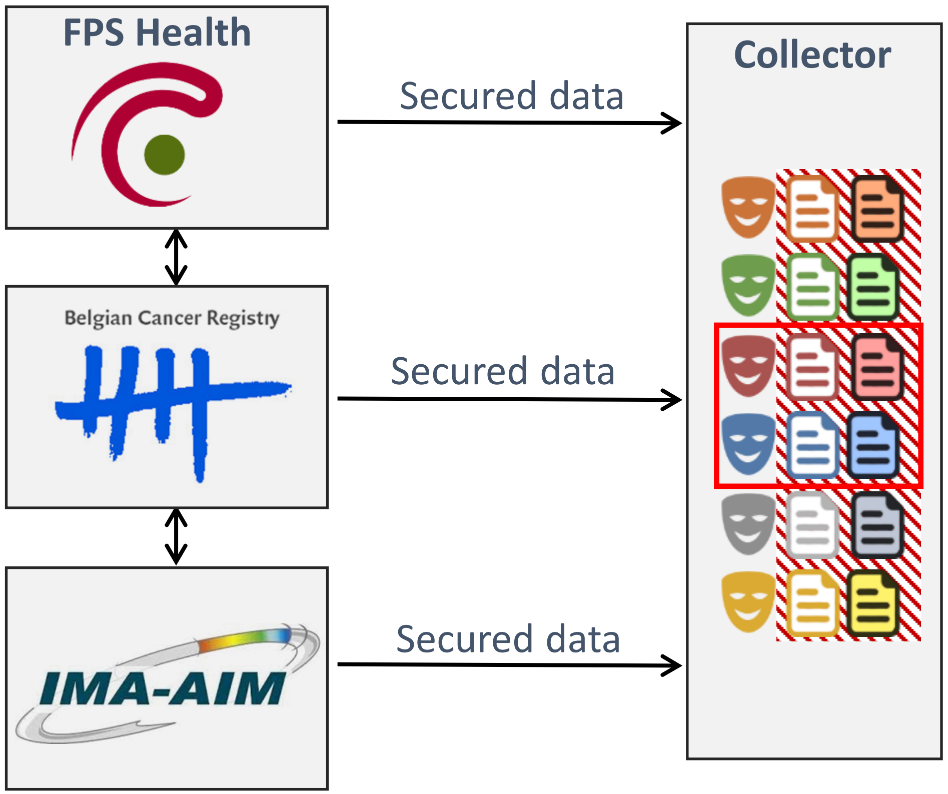


Extensible from there

Oblivious Join

- Problem statement
- **Concept**
- In practice
- Conclusion

Concept



Data sources

- ❖ Do not learn any new personal or statistical data
- ❖ Only see identifiers of their data

Collector

- ❖ Learns only minimum required pseudonymised personal data
- ❖ Learns high-level statistical data
  E.g. number of citizens with cancer diagnosis
- ❖ Only sees pseudonyms

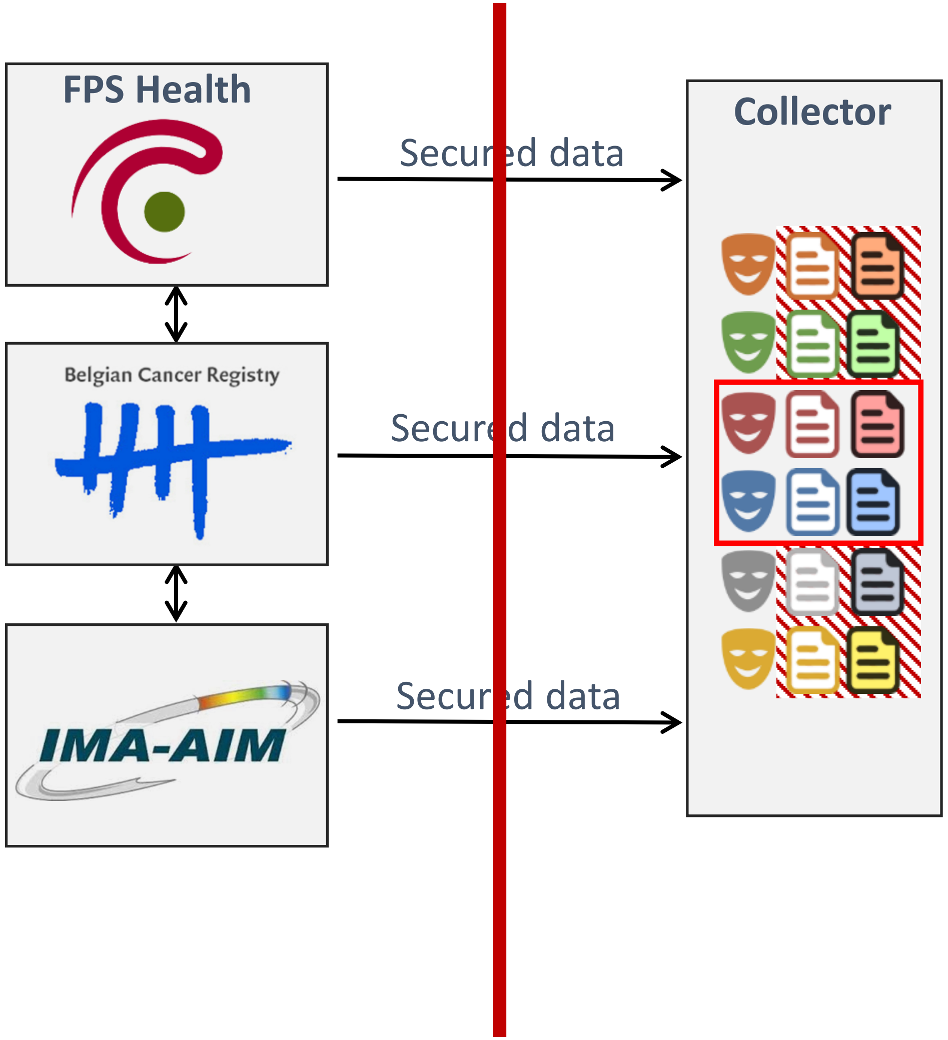
Properties

- ✓ Privacy-friendly & secure
- ✓ Distributed: no pseudon. service
- ✓ Uniform & no integration
- ✓ Fast & cost-efficient

3 steps protocol

1. Fully automated agreements between data sources (no human intervention)
2. Each data source sends all potentially relevant data encrypted & pseudonymised to collector
3. Thanks previous agreements (step 1) collector can only decrypt & combine pertinent records

Concept



No collusion between data source and collector

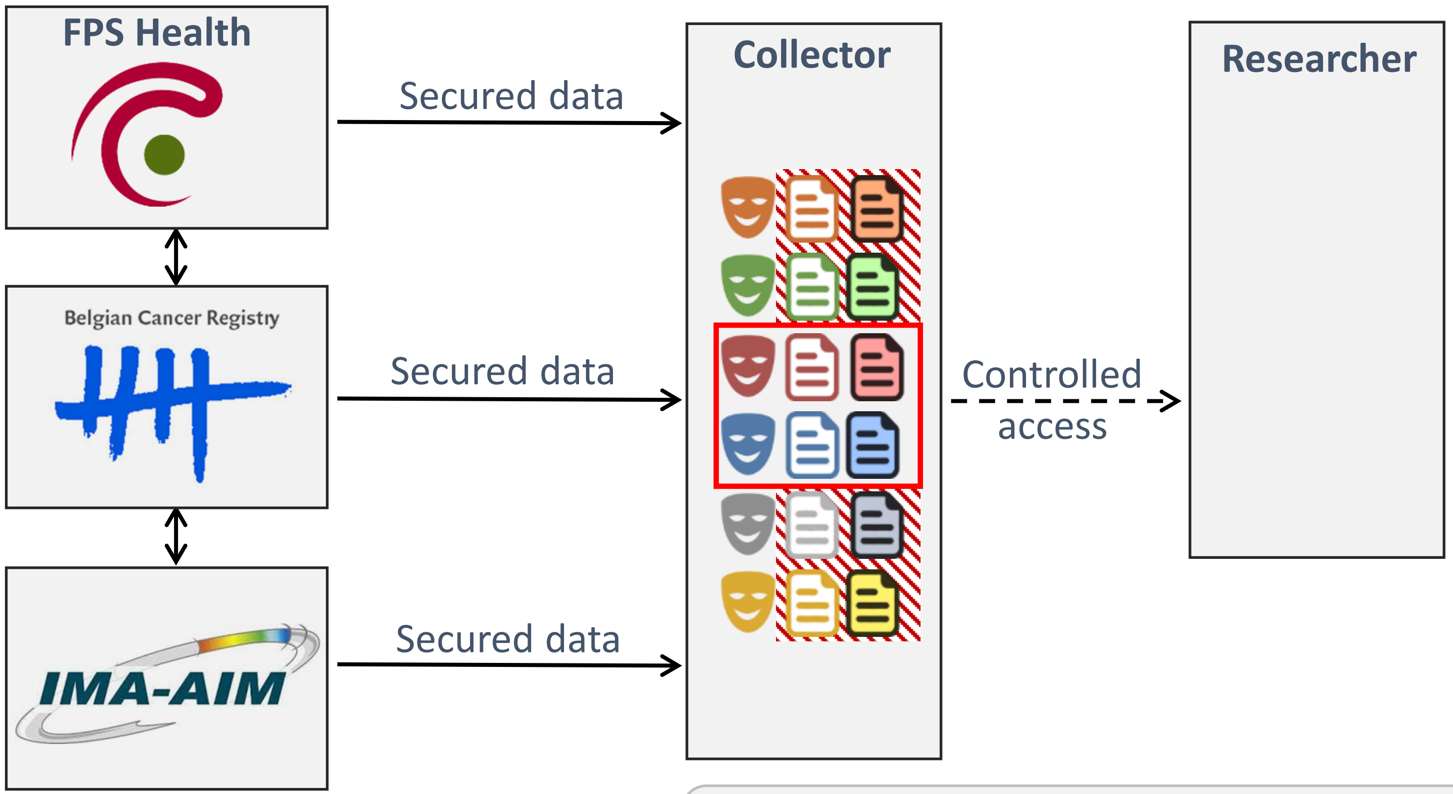
Properties

- ✓ Privacy-friendly & secure
- ✓ Distributed: no pseudon. service
- ✓ Uniform & no integration
- ✓ Fast & cost-efficient

3 steps protocol

1. Fully automated agreements between data sources (no human intervention)
2. Each data source sends all potentially relevant data encrypted & pseudonymised to collector
3. Thanks previous agreements (step 1) collector can only decrypt & combine pertinent records

Concept



Collector

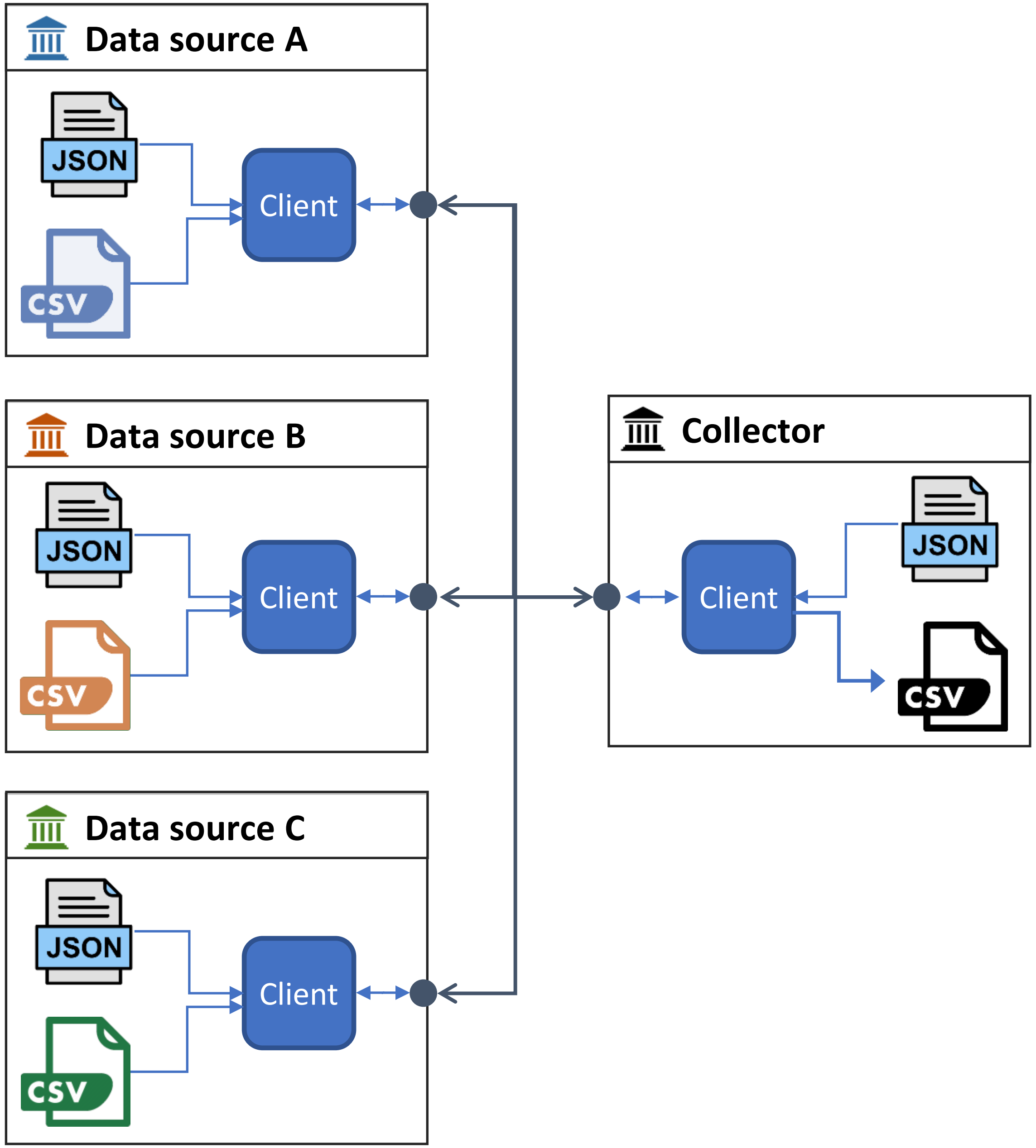
Independent and semi-trusted

1. Deletes asap irrelevant ciphertexts
2. Can do additional checks on the data
3. Controlled access to researcher

Oblivious Join

- Problem statement
- Concept
- **In practice**
- Conclusion

In practice



client

Client

- Java jar
- No integration required → non-intrusive, flexible
- All parties use same client (software)
- Command-line interface

JSON

Project description

- JSON file
- Created by coordinating party
- Contains all info required to execute protocol
- All parties use same project description

CSV

Input files

- CSV file
- Created by individual data source (out of scope)
- Contains all, potentially relevant, identified personal data

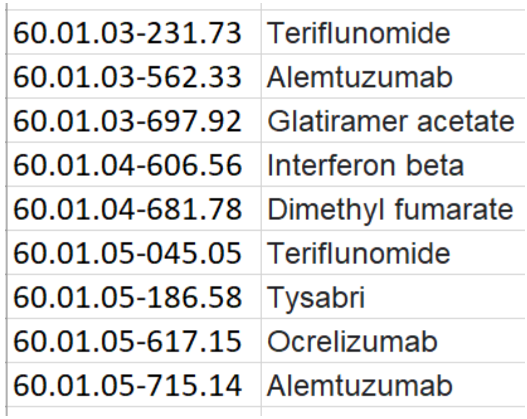
CSV

Output file

- CSV file
- Collector’s output after protocol execution
- Contains minimal required joined & pseudonymised personal data

Test with fictional data

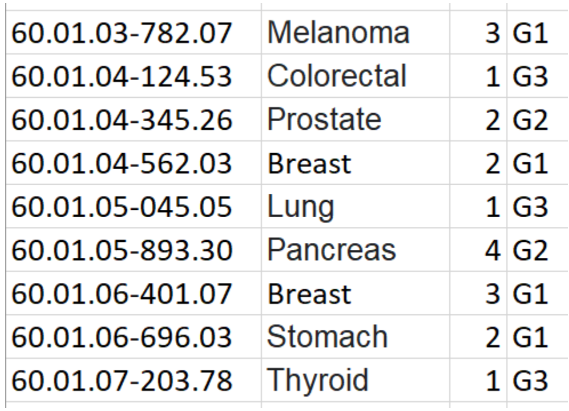
Extract input CSV
Data source 1 (IMA-AIM)



60.01.03-231.73	Teriflunomide
60.01.03-562.33	Alemtuzumab
60.01.03-697.92	Glatiramer acetate
60.01.04-606.56	Interferon beta
60.01.04-681.78	Dimethyl fumarate
60.01.05-045.05	Teriflunomide
60.01.05-186.58	Tysabri
60.01.05-617.15	Ocrelizumab
60.01.05-715.14	Alemtuzumab

200 000 records
E.g. Citizens with MS

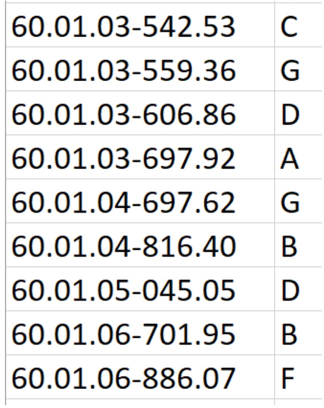
Extract input CSV
Data source 2 (BCR)



60.01.03-782.07	Melanoma	3	G1
60.01.04-124.53	Colorectal	1	G3
60.01.04-345.26	Prostate	2	G2
60.01.04-562.03	Breast	2	G1
60.01.05-045.05	Lung	1	G3
60.01.05-893.30	Pancreas	4	G2
60.01.06-401.07	Breast	3	G1
60.01.06-696.03	Stomach	2	G1
60.01.07-203.78	Thyroid	1	G3

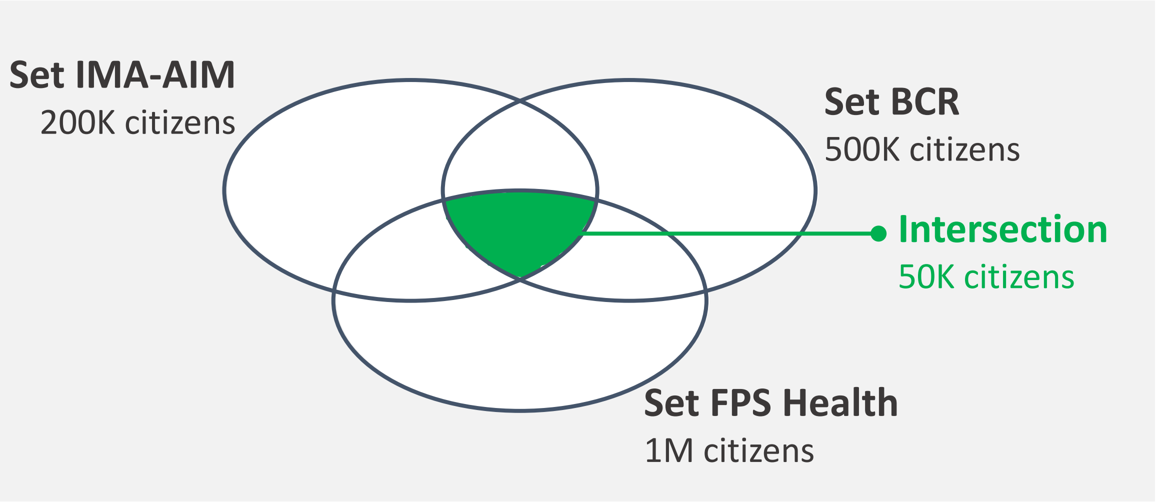
500 000 records
E.g. Citizens with cancer

Extract input CSV
Data source 3 (FPS Health)



60.01.03-542.53	C
60.01.03-559.36	G
60.01.03-606.86	D
60.01.03-697.92	A
60.01.04-697.62	G
60.01.04-816.40	B
60.01.05-045.05	D
60.01.06-701.95	B
60.01.06-886.07	F

1 000 000 records
E.g. Citizens with high-risk profile



Set IMA-AIM
200K citizens

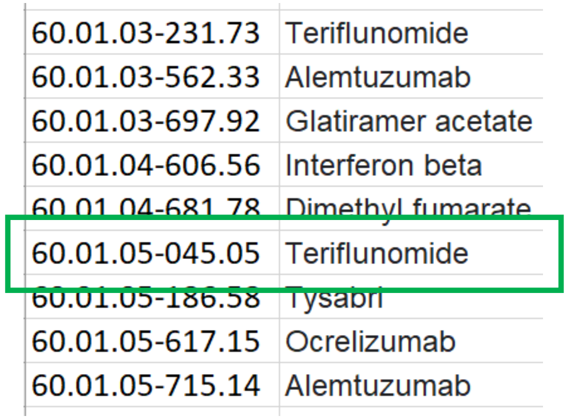
Set BCR
500K citizens

Intersection
50K citizens

Set FPS Health
1M citizens

Test with fictional data

Extract input CSV
Data source 1 (IMA-AIM)

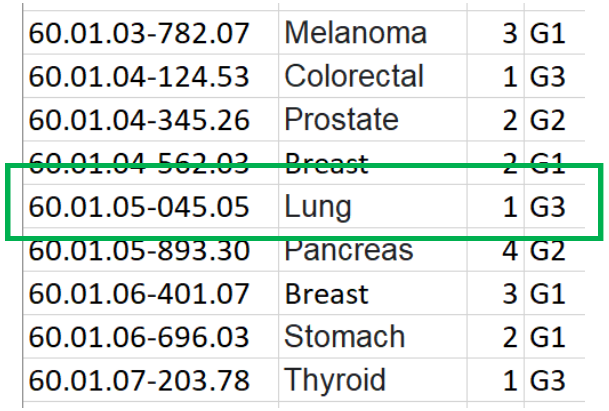


60.01.03-231.73	Teriflunomide
60.01.03-562.33	Alemtuzumab
60.01.03-697.92	Glatiramer acetate
60.01.04-606.56	Interferon beta
60.01.04-681.78	Dimethyl fumarate
60.01.05-045.05	Teriflunomide
60.01.05-186.58	Tysabri
60.01.05-617.15	Ocrelizumab
60.01.05-715.14	Alemtuzumab

200 000 records

E.g. Citizens with MS

Extract input CSV
Data source 2 (BCR)

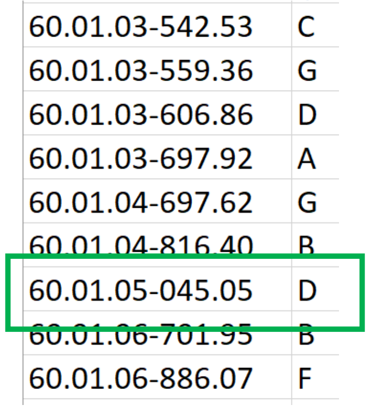


60.01.03-782.07	Melanoma	3	G1
60.01.04-124.53	Colorectal	1	G3
60.01.04-345.26	Prostate	2	G2
60.01.04-562.03	Breast	2	G1
60.01.05-045.05	Lung	1	G3
60.01.05-893.30	Pancreas	4	G2
60.01.06-401.07	Breast	3	G1
60.01.06-696.03	Stomach	2	G1
60.01.07-203.78	Thyroid	1	G3

500 000 records

E.g. Citizens with cancer

Extract input CSV
Data source 3 (FPS Health)



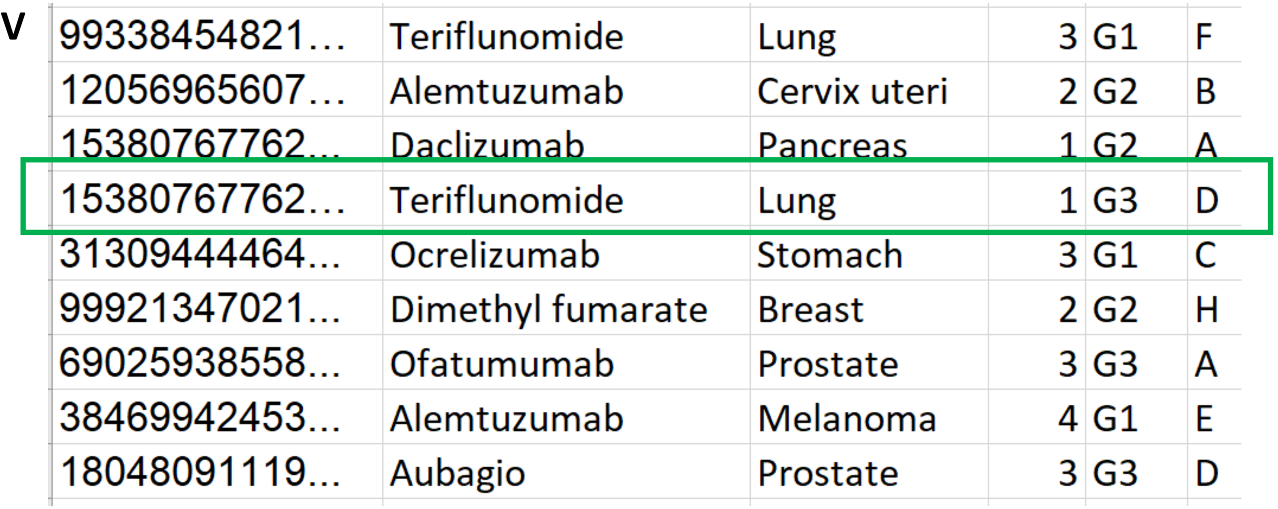
60.01.03-542.53	C
60.01.03-559.36	G
60.01.03-606.86	D
60.01.03-697.92	A
60.01.04-697.62	G
60.01.04-816.40	B
60.01.05-045.05	D
60.01.06-701.95	B
60.01.06-886.07	F

1 000 000 records

E.g. Citizens with high-risk profile

Extract output CSV
Collector (KSZ)

50 000 records



99338454821…	Teriflunomide	Lung	3	G1	F
12056965607…	Alemtuzumab	Cervix uteri	2	G2	B
15380767762…	Daclizumab	Pancreas	1	G2	A
15380767762…	Teriflunomide	Lung	1	G3	D
31309444464…	Ocrelizumab	Stomach	3	G1	C
99921347021…	Dimethyl fumarate	Breast	2	G2	H
69025938558…	Ofatumumab	Prostate	3	G3	A
38469942453…	Alemtuzumab	Melanoma	4	G1	E
18048091119…	Aubagio	Prostate	3	G3	D

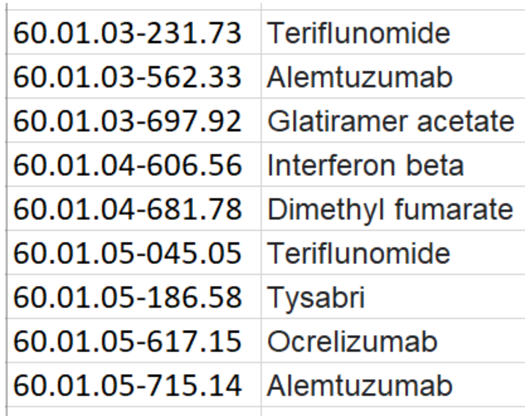
Who sees what?

- Data sources only see identifiers
- Collector only sees pseudonyms
- No pseudonymisation service

Test with fictional data

Extract input CSV

Data source 1 (IMA-AIM)



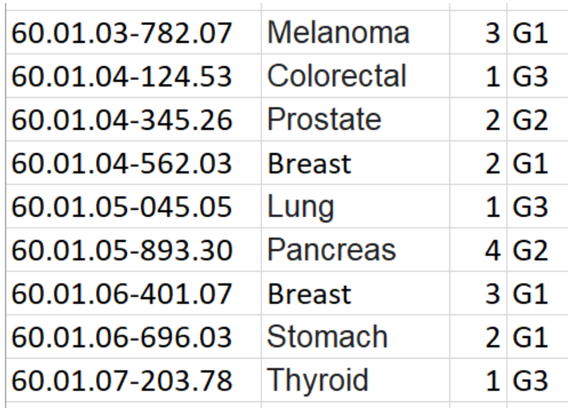
60.01.03-231.73	Teriflunomide
60.01.03-562.33	Alemtuzumab
60.01.03-697.92	Glatiramer acetate
60.01.04-606.56	Interferon beta
60.01.04-681.78	Dimethyl fumarate
60.01.05-045.05	Teriflunomide
60.01.05-186.58	Tysabri
60.01.05-617.15	Ocrelizumab
60.01.05-715.14	Alemtuzumab

200 000 records

E.g. Citizens with MS

Extract input CSV

Data source 2 (BCR)



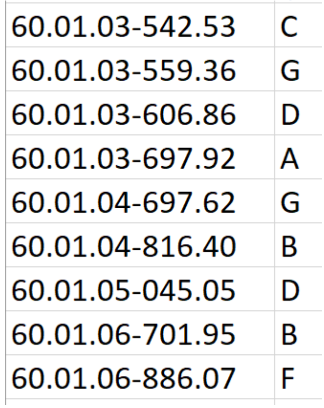
60.01.03-782.07	Melanoma	3	G1
60.01.04-124.53	Colorectal	1	G3
60.01.04-345.26	Prostate	2	G2
60.01.04-562.03	Breast	2	G1
60.01.05-045.05	Lung	1	G3
60.01.05-893.30	Pancreas	4	G2
60.01.06-401.07	Breast	3	G1
60.01.06-696.03	Stomach	2	G1
60.01.07-203.78	Thyroid	1	G3

500 000 records

E.g. Citizens with cancer

Extract input CSV

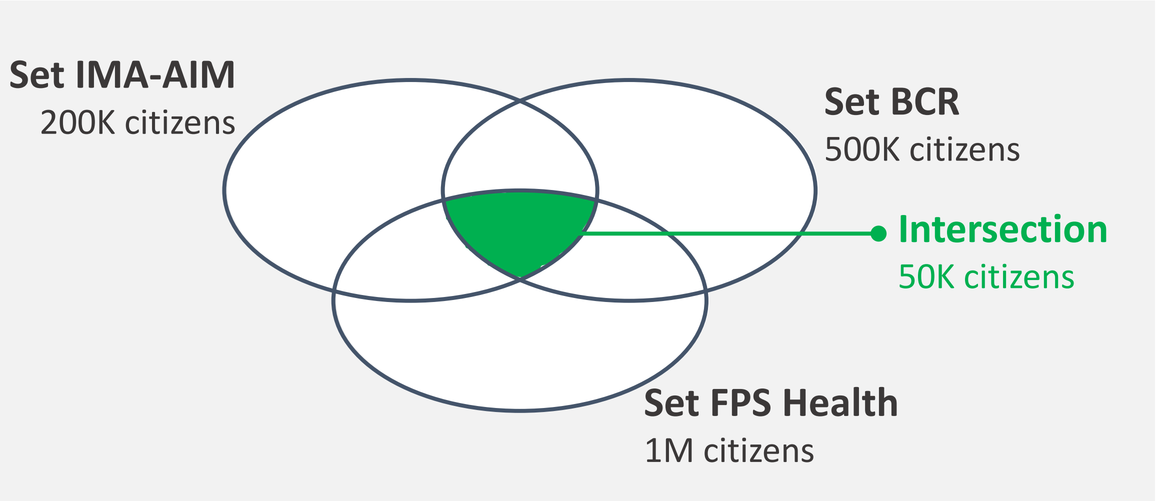
Data source 3 (FPS Health)



60.01.03-542.53	C
60.01.03-559.36	G
60.01.03-606.86	D
60.01.03-697.92	A
60.01.04-697.62	G
60.01.04-816.40	B
60.01.05-045.05	D
60.01.06-701.95	B
60.01.06-886.07	F

1 000 000 records

E.g. Citizens with high-risk profile



Performance test

Parameters

- MinNbRecords: 10
- 128 bit security

Infrastructure

- Data sources: 4 i9-7940x cores @ 3.10 GHz, 16GB RAM
- Collector: 2 i9-7940x cores @ 3.10 GHz , 16GB RAM

Results

- **< 2 min calculations**
- Excl. a few hundred MBs data transfer

Oblivious Join

- Problem statement
- Concept
- In practice
- **Conclusion**

Collaboration universities

Interdisciplinary paper (To be published in 2024)

Privacy-By-Design in the Belgian Public Sector

Pseudonymising & Joining Personal Data Fragmented over Multiple Organisations

KU LEUVEN

Smals ICT for society

In *Public Governance and Emerging Technologies – Values, Trust, and Compliance by Design*

CHAIN Connecting Humankind through Algorithms and Information Networks

Utrecht University

SPRINGER NATURE

Expert paper

Ongoing
<CONFIDENTIAL>

https://chainresearch.eu/
https://www.uu.nl/en/events/conference-public-governance-and-emerging-technologies-values-trust-and-compliance-by-design

Evaluation

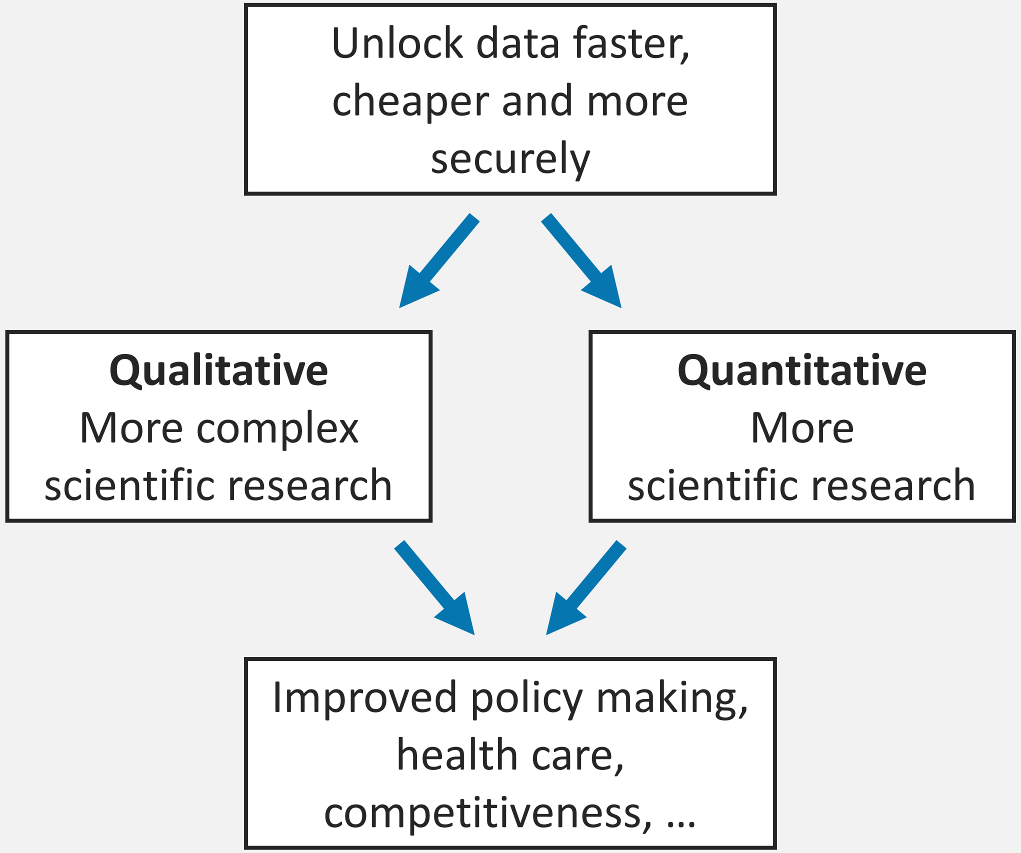
Advantages

- ✓ Answer on business need
- ✓ Privacy-friendly & secure
- ✓ Distributed (no pseudonymisation service)
- ✓ Uniform & no integration
- ✓ Fast & cost-efficient
- ✓ Formal academic validation

Challenges

- Only passive interest
- Still in research phase
- Higher development complexity (but lower infra)
- Extensions required

Opportunities



Unlock data faster, cheaper and more securely

Qualitative
More complex scientific research

Quantitative
More scientific research

Improved policy making, health care, competitiveness, …

Wrapping up

Innovation @ Smals Research

Smart Pseudonymisation

Conversion from citizen identifiers to pseudonyms

Format-Preserving Pseudonymisation

Retroactive protection of personal data in TEST & ACC of legacy applications

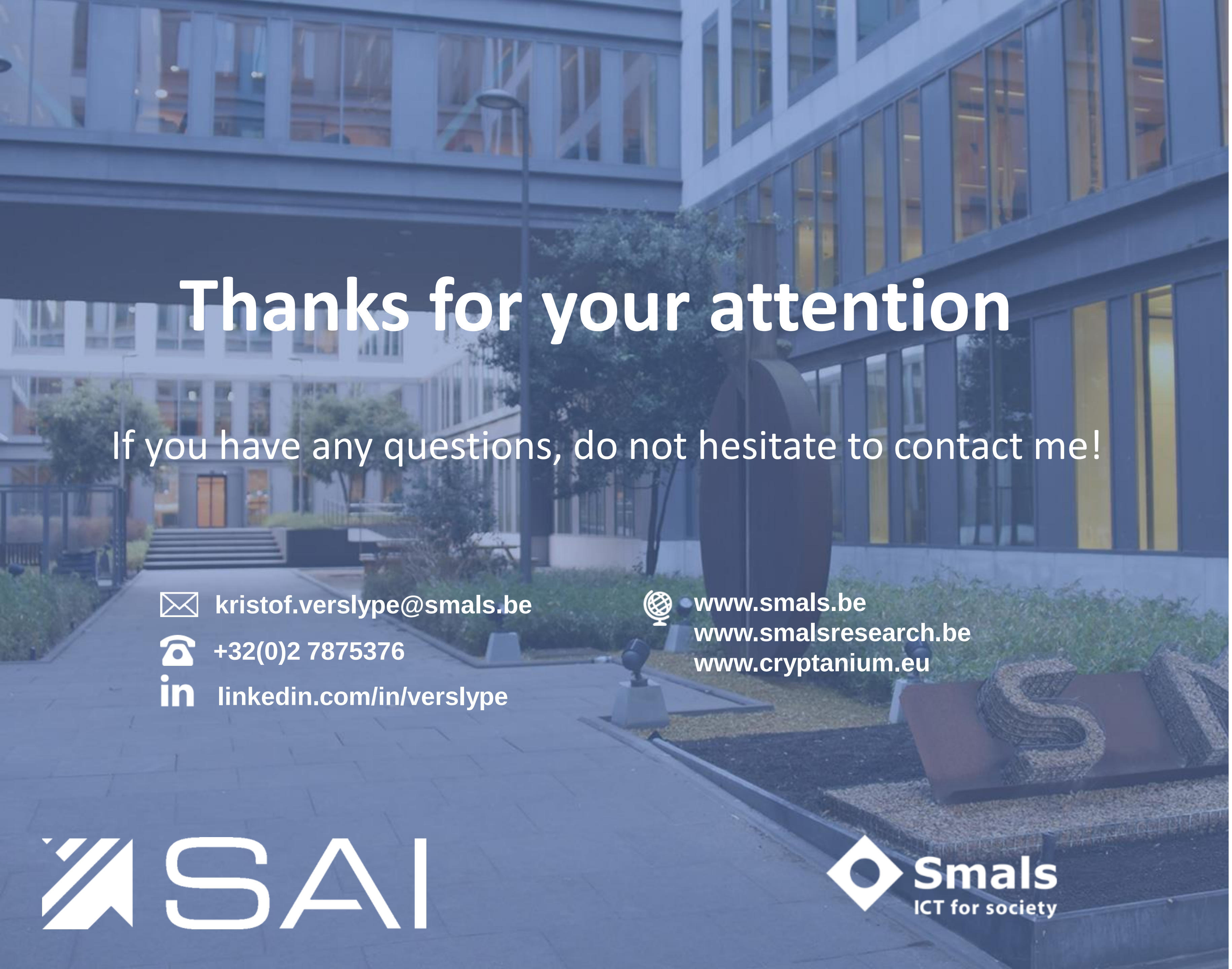
eHealth Blind Pseudonymisation

Proactive protection of personal data in applications
Privacy by Design

Oblivious Join

Non-trivial join & pseudonymise projects for research purposes
Distributed & no integration

Smart pseudonymisation can play a crucial role to protect personal data



Thanks for your attention

If you have any questions, do not hesitate to contact me!

✉ kristof.verslype@smals.be

☎ +32(0)2 7875376

in linkedin.com/in/verslype

www.smals.be
www.smalsresearch.be
www.cryptanium.eu

SAI

Smals
ICT for society

Further reading
www.smalsresearch.be

www.smalsresearch.be/tag/pseudonymisation/

Images

Judy Dean

Creative Commons

https://flickr.com/photos/peterscherub/53152339550/

Pixabay

Pixabay License

https://pixabay.com/fr/photos/femme-les-yeux-masquer-carnaval-411494/

Daniel Bruce

Creative Commons

https://iconscout.com/free-icon/mask-126

estorde

Creative commons

https://flickr.com/photos/estorde/4572006561

Aris Gionis

Creative Commons

Flickr

Oscar Gende Villar

Creative Commons

Flickr