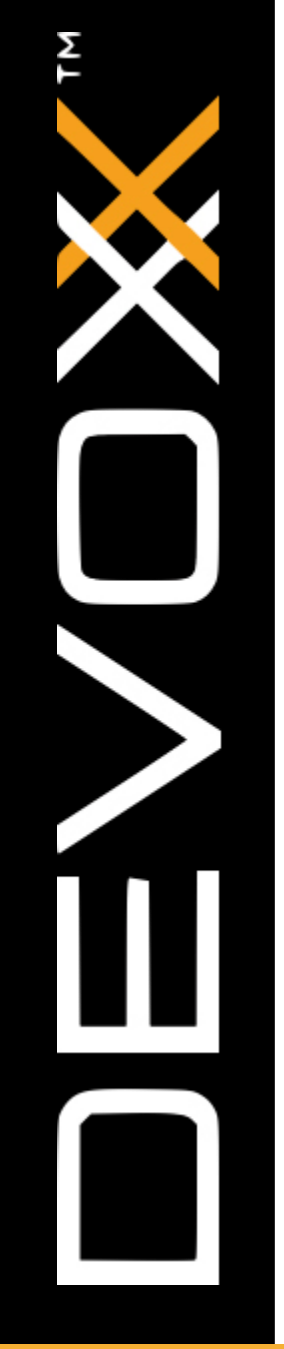
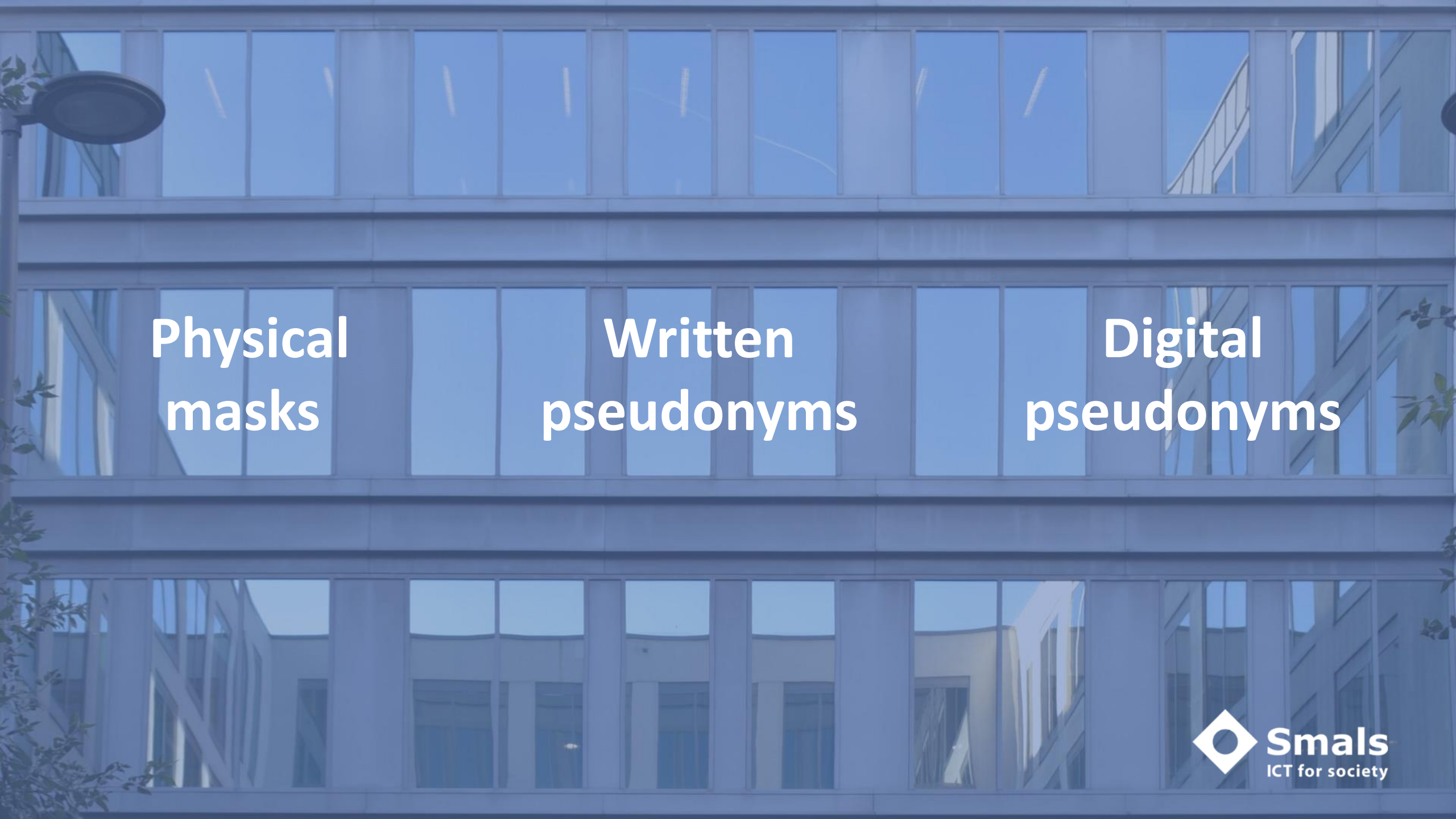
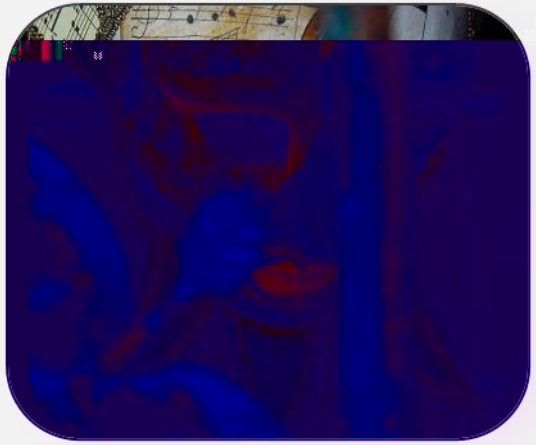

DEVOXX
TM


Physical masks
Written pseudonyms
Digital pseudonyms
Smals
ICT for society

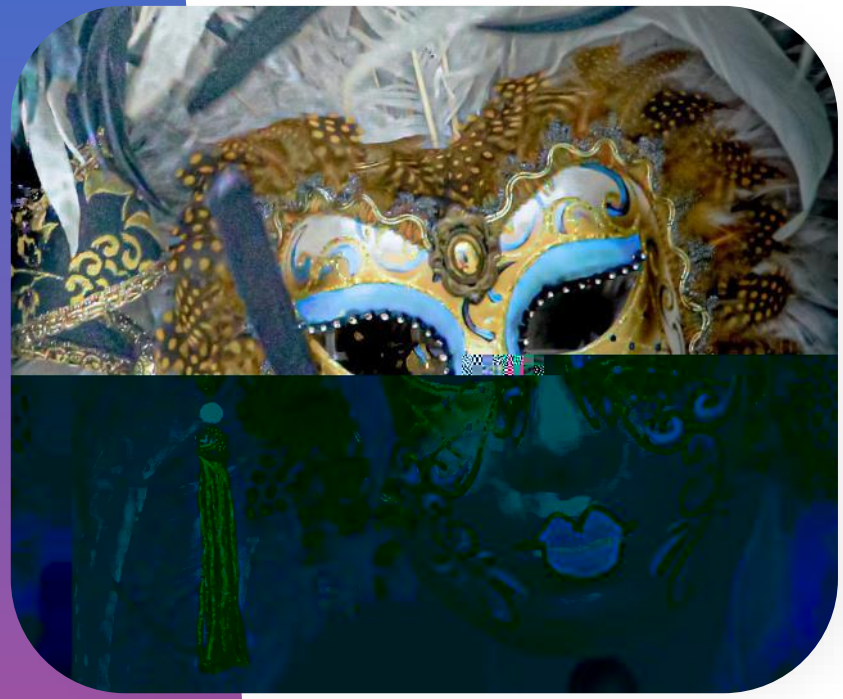
Innovation @ Smals Research

Smart Pseudonymisation



ICT for society

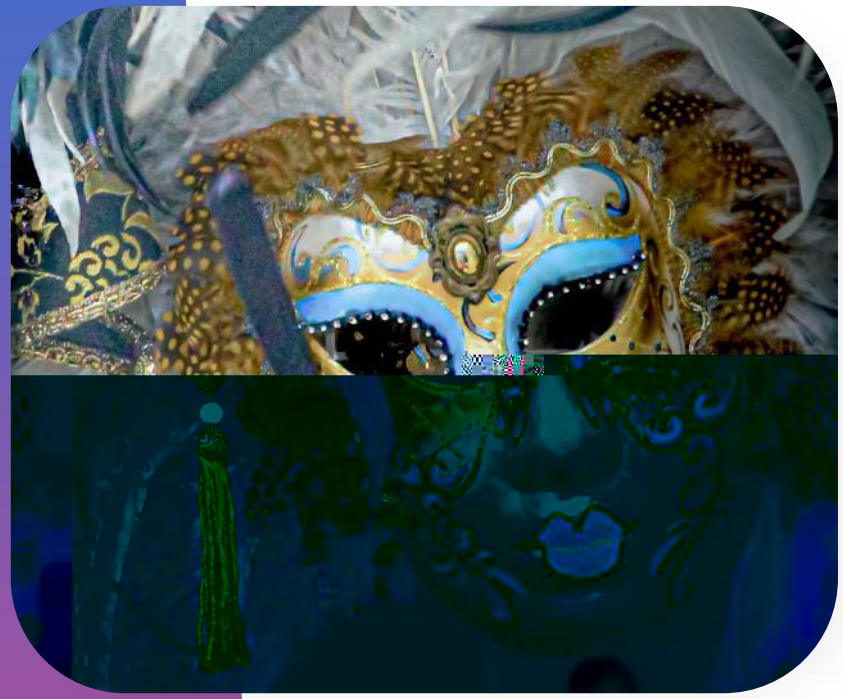
Format-Preserving Pseudonymisation



- Problem statement
- Concept
- Experimental service
- Conclusion

Smals
ICT for society

Format-Preserving Pseudonymisation



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

Smals
ICT for society

Widespread use of personal data in non-prod environments

Security

2016

2021

2022

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

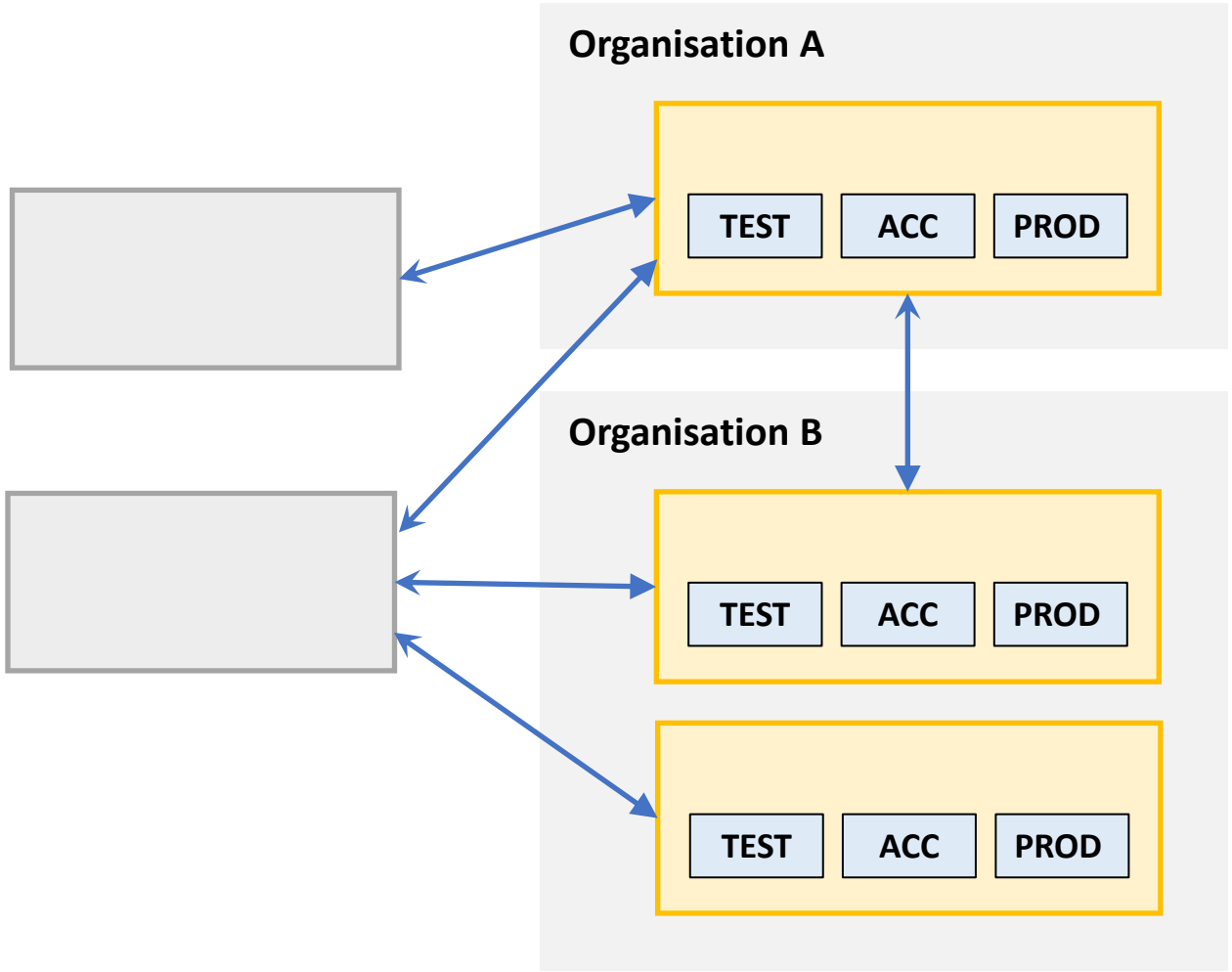
GDPR, Art 32.
[…] *the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:*

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

b) […]

Smals
ICT for society

Reality in public sector



Question customer

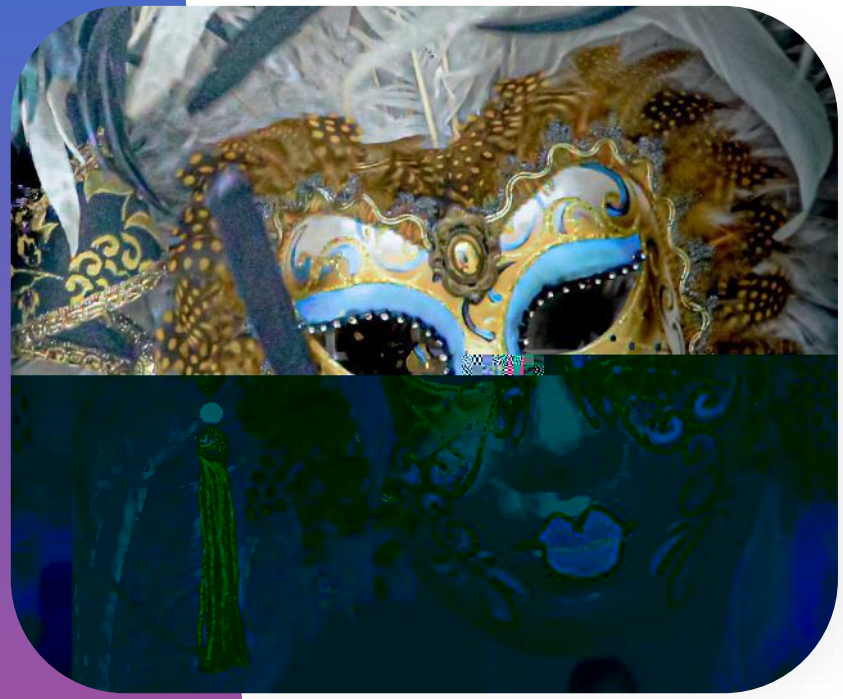
How to improve privacy in TEST & ACC?

not

Because

Smals
ICT for society

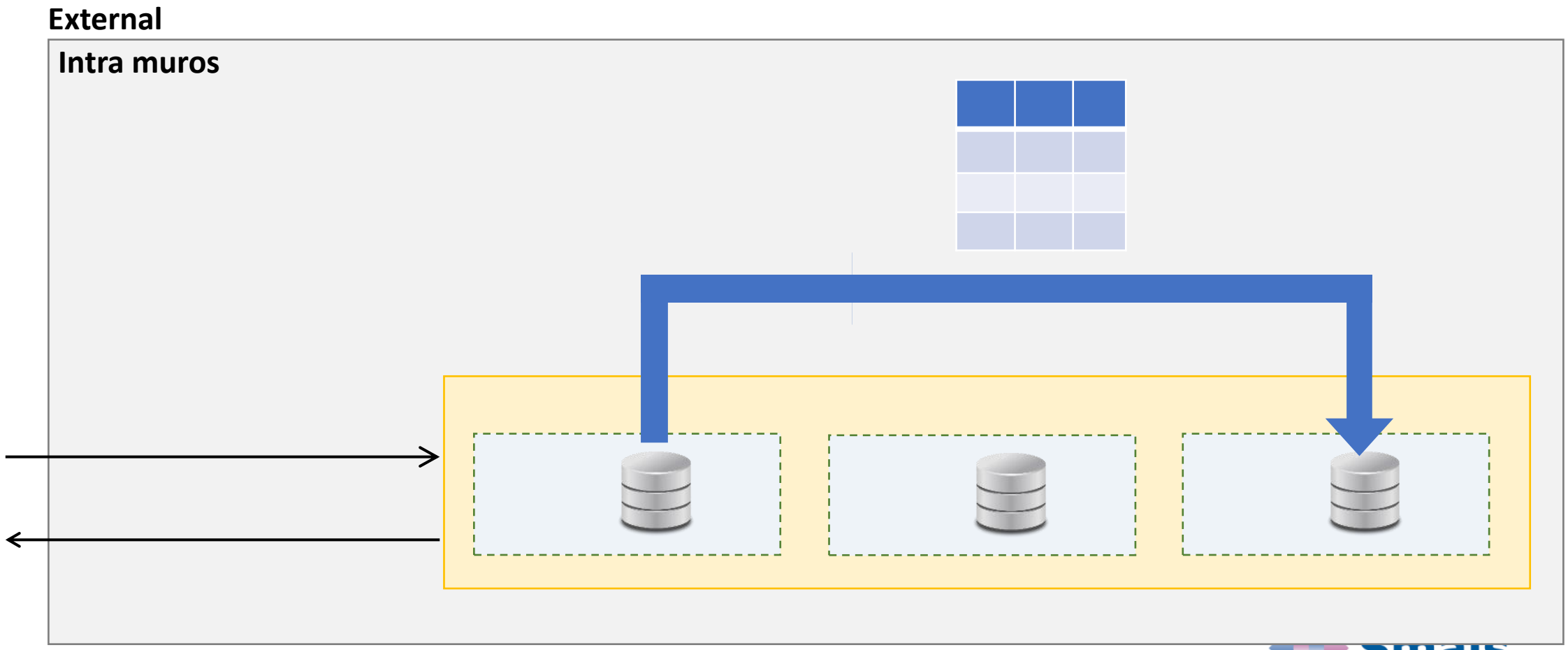
Format-Preserving Pseudonymisation



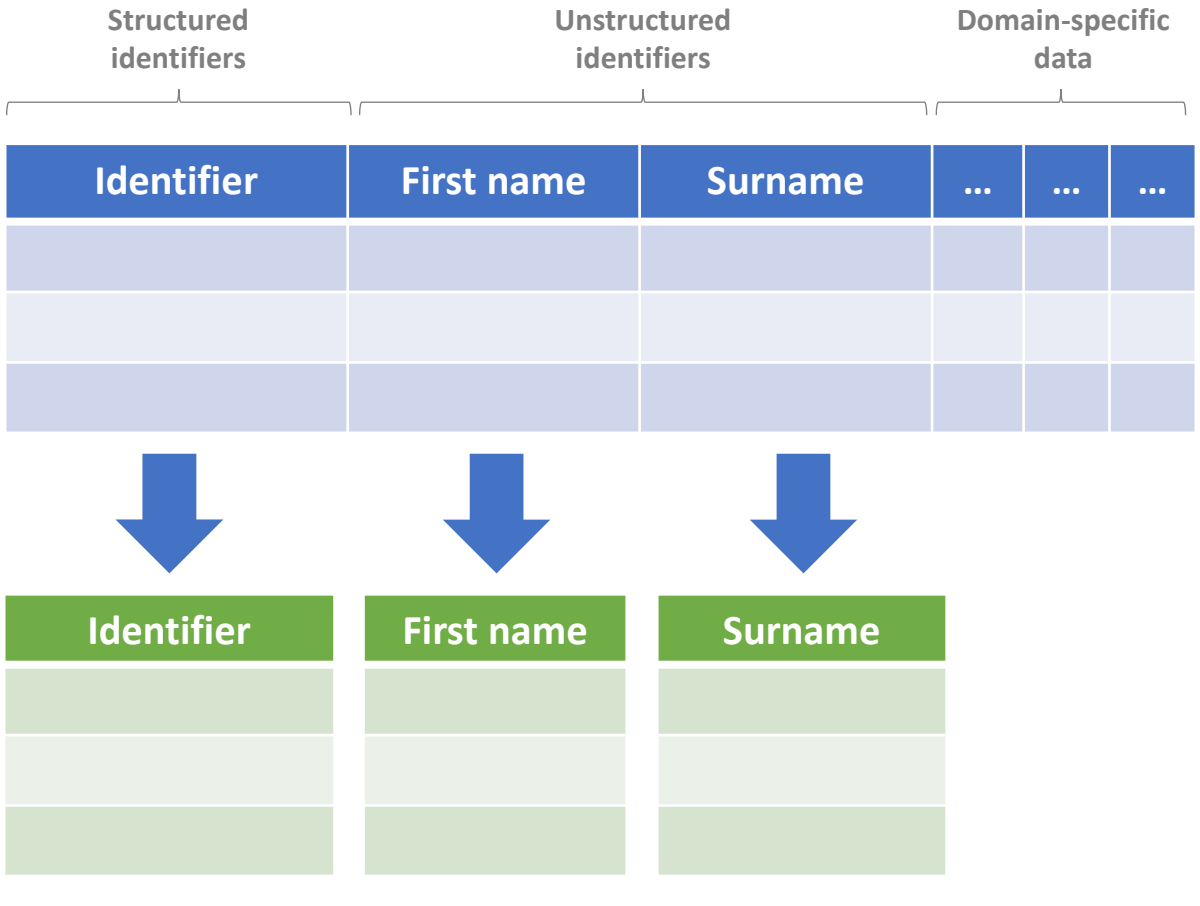
- Problem statement
- **Concept**
- Experimental service
- Conclusion

Smals
ICT for society

Current practice



Approach by member



1. **Pseudonymise**
   - Bidirectional
   - By Smals Research
2. **Shuffle**
   - Unidirectional
   - By Customer

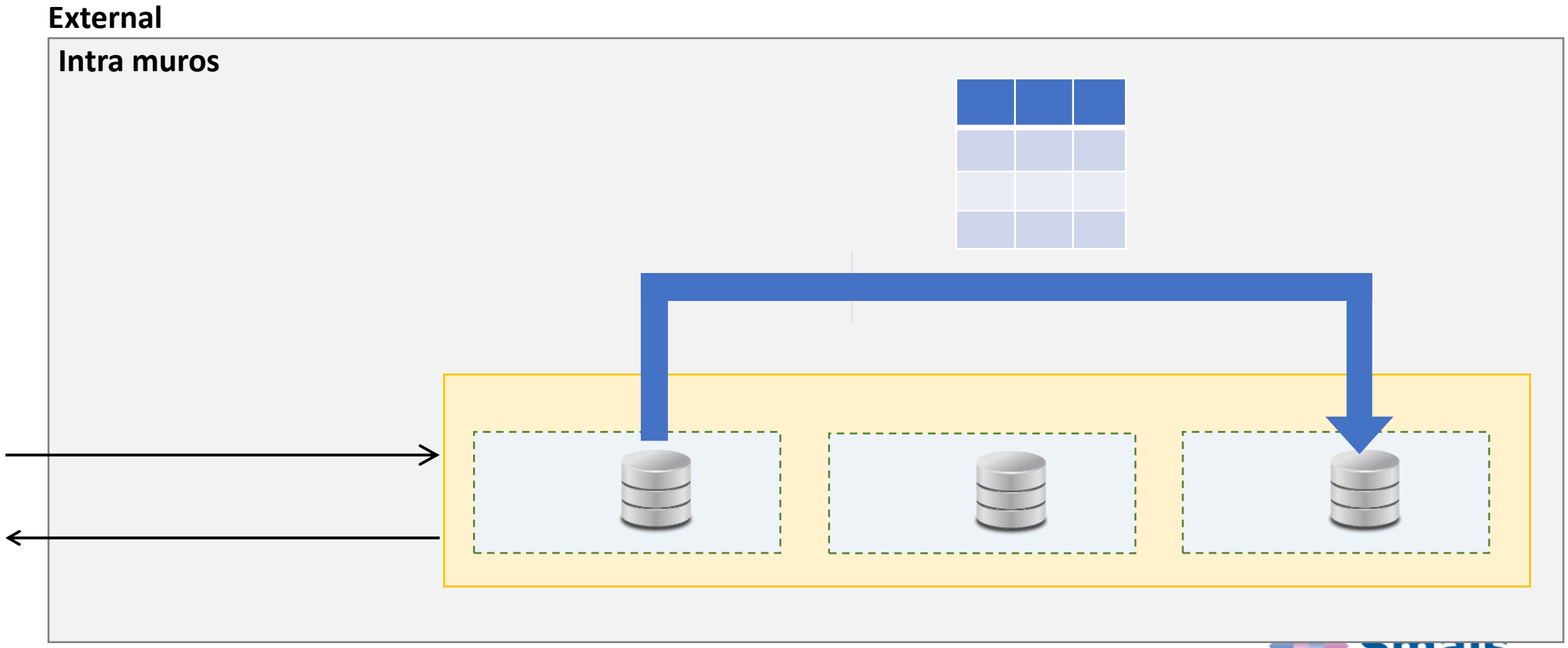
Transformed snapshot

Identifier	First name	Surname	...	...	...

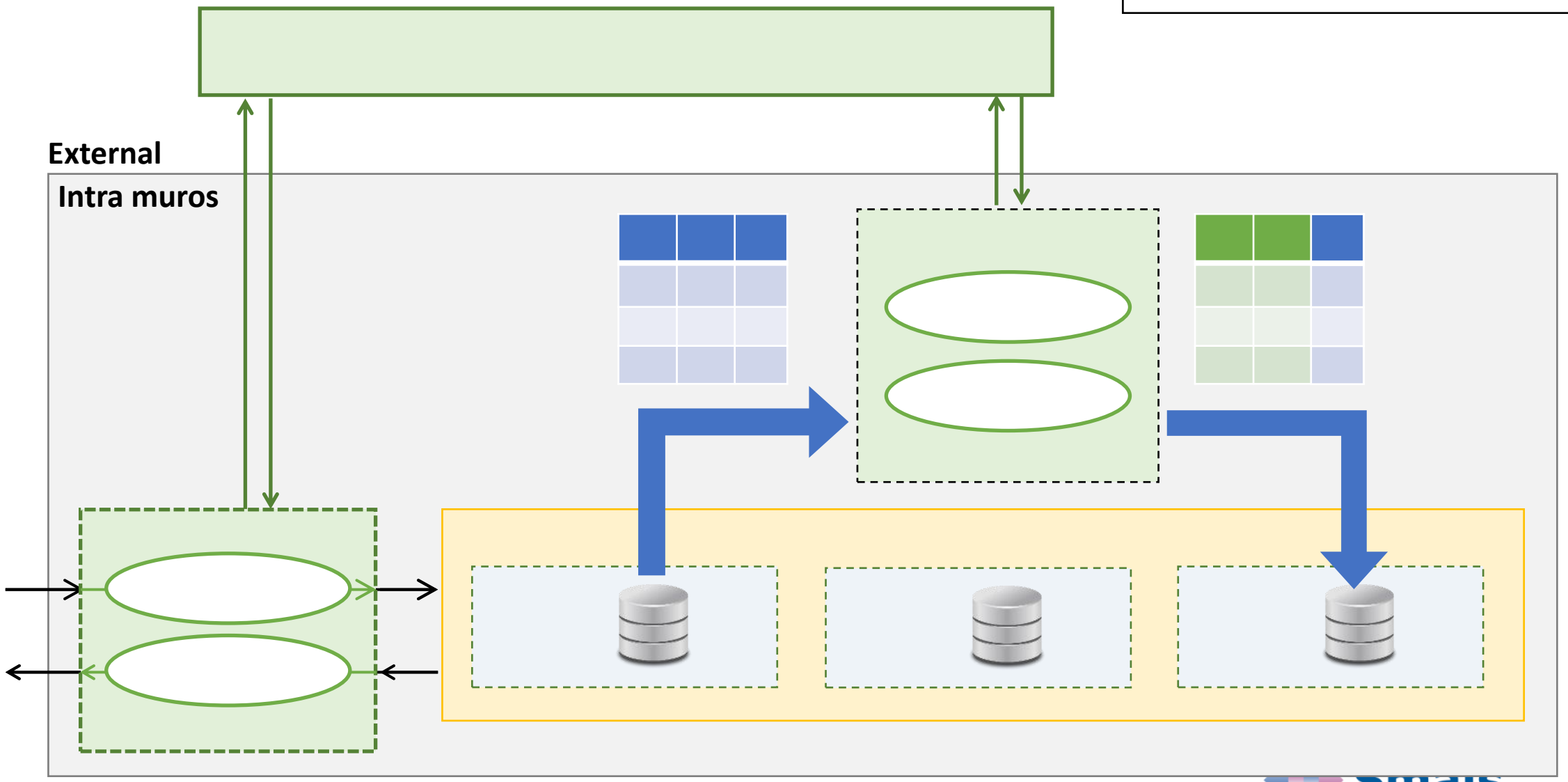
Records useful for TEST & ACC, while hard to identify!

Smals
ICT for society

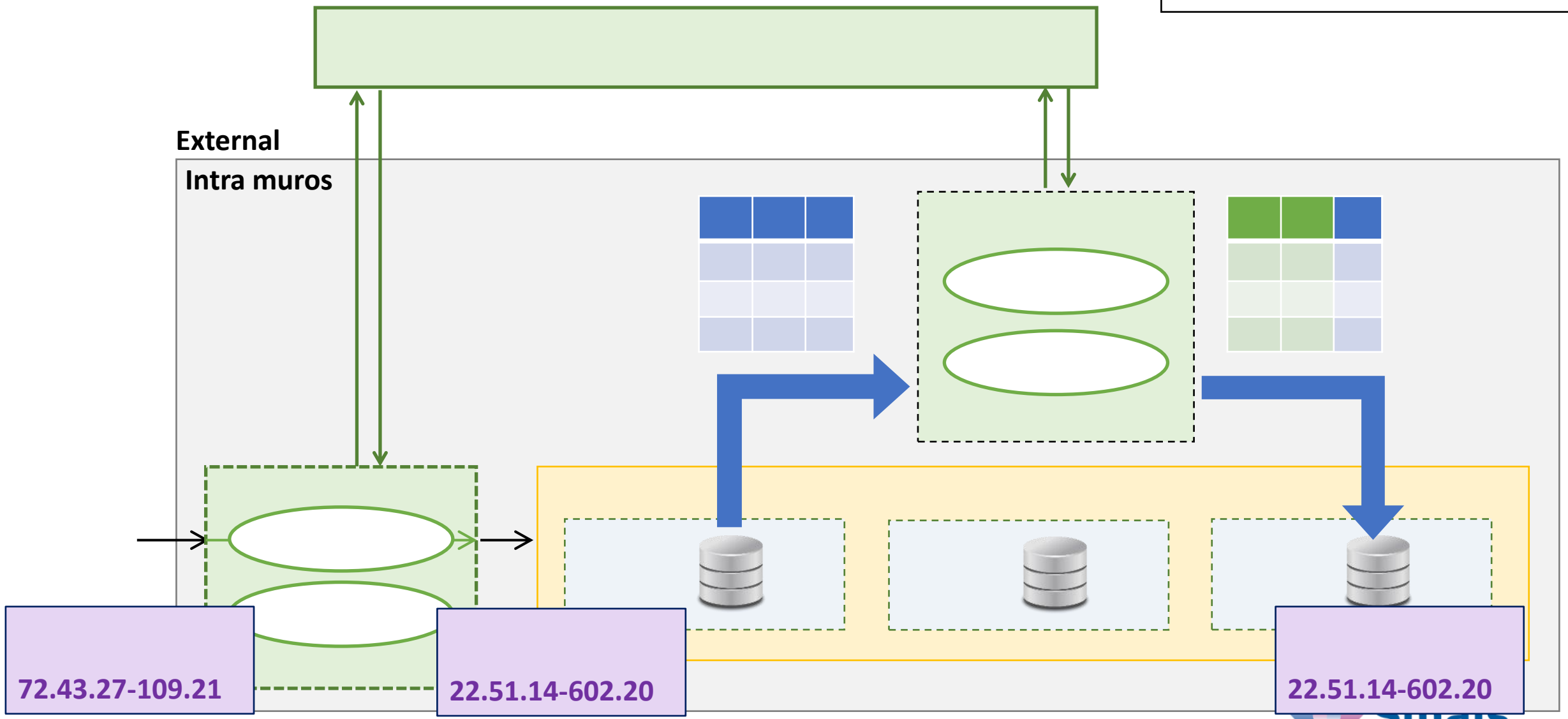
PoC in collaboration with customer



PoC in collaboration with customer



PoC in collaboration with customer

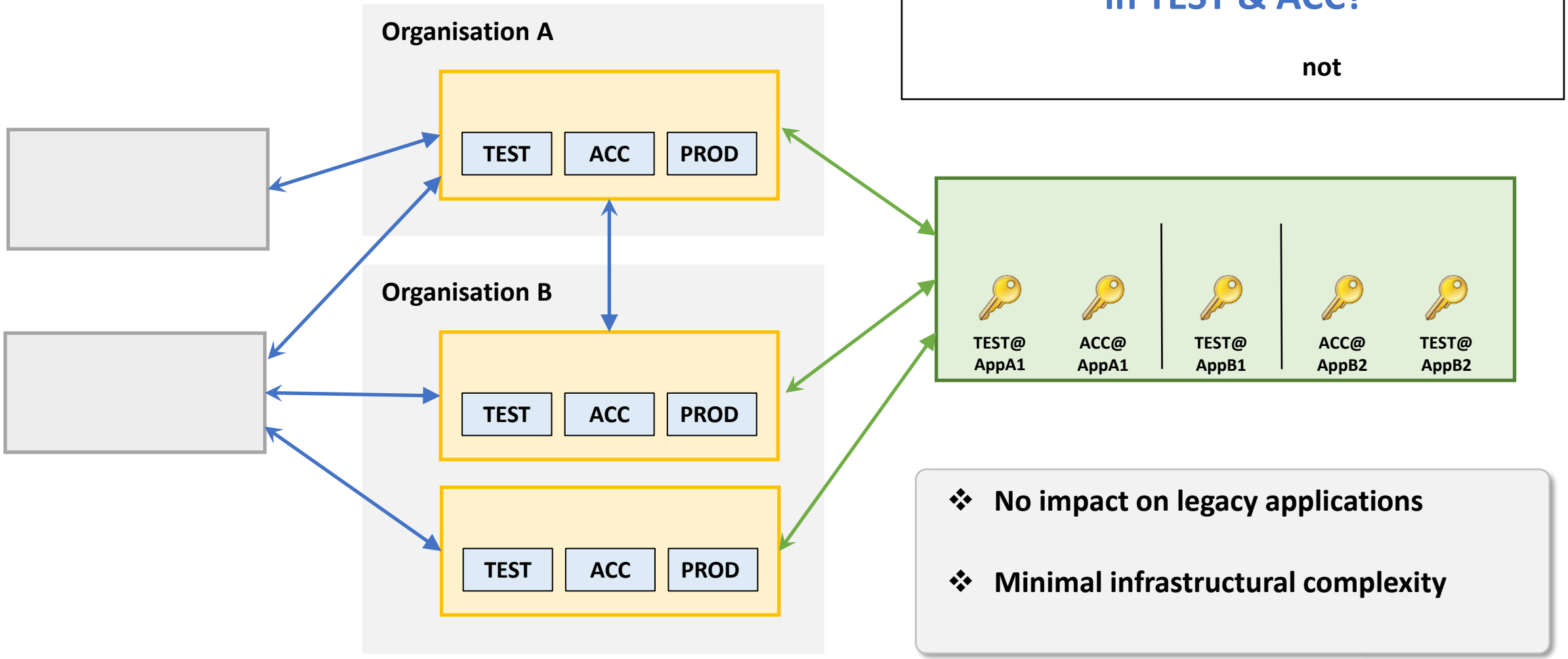


Reality in public sector

Question customer

How to improve privacy in TEST & ACC?

not

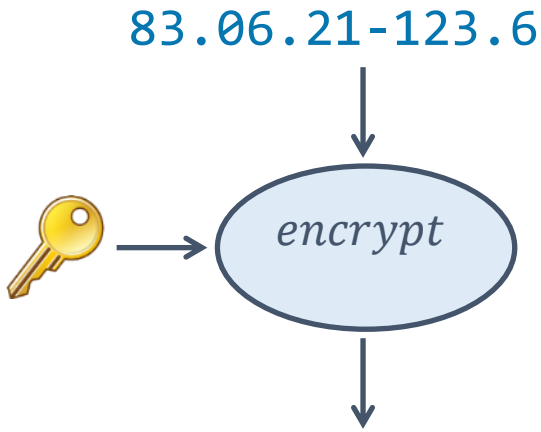


- ❖ **No impact on legacy applications**
- ❖ **Minimal infrastructural complexity**

Smals ICT for society

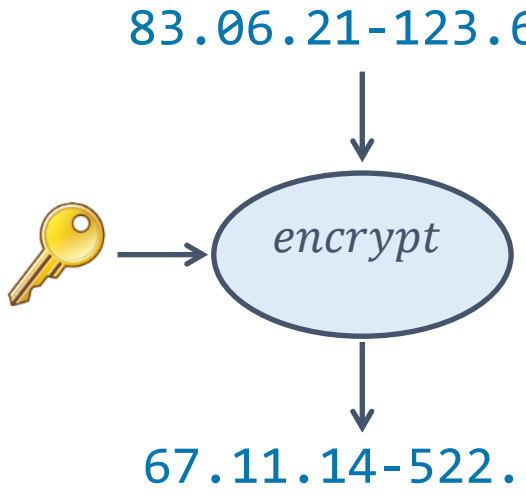
Encryption

Traditional encryption



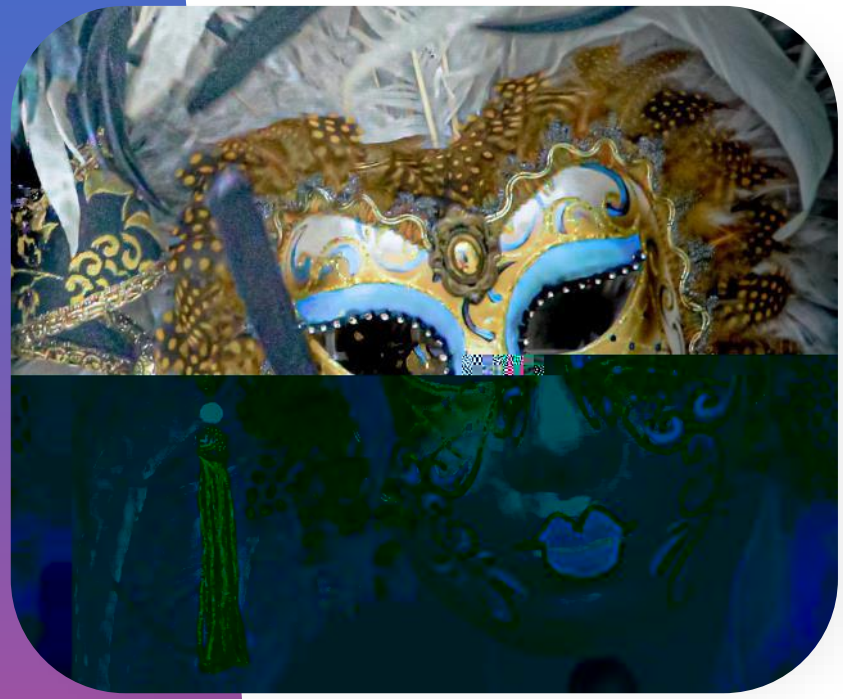

```
70 58 e5 87 1b 9e fc d2
0e a5 54 87 ce b6 34 e7
c3 66 ec 08 0e 2f 0c 2c
51 a8 77 11 28 44 5c e1
34 ff 95 d5 3e 44 75 c3
ed 2b 0a 58 3a 67 74 dd
ec 1e ba e0 3a 3c c5 7f
75 59 c2 53 4a 73 a2 d9
```

Format-preserving encryption



-
-
-

Format-Preserving Pseudonymisation



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

Smals
ICT for society

Experimental REST service

B M M

Rest API

✓ M BM B

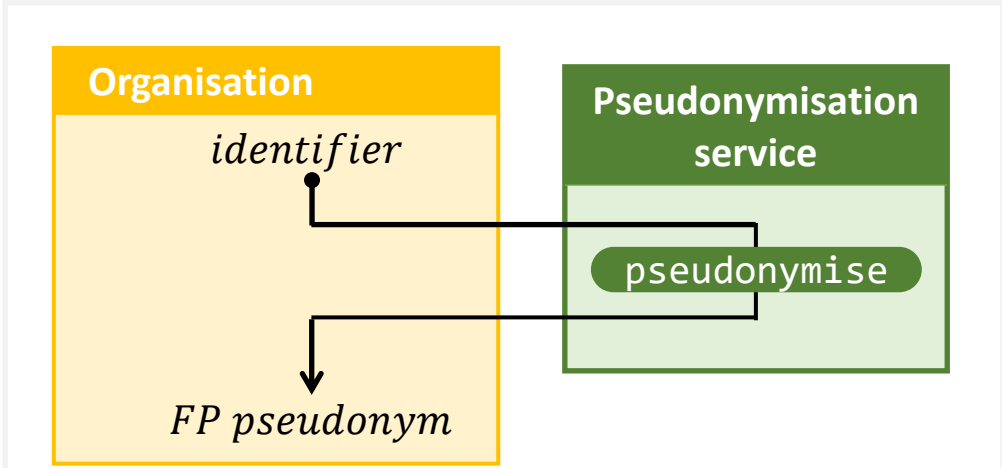
✓

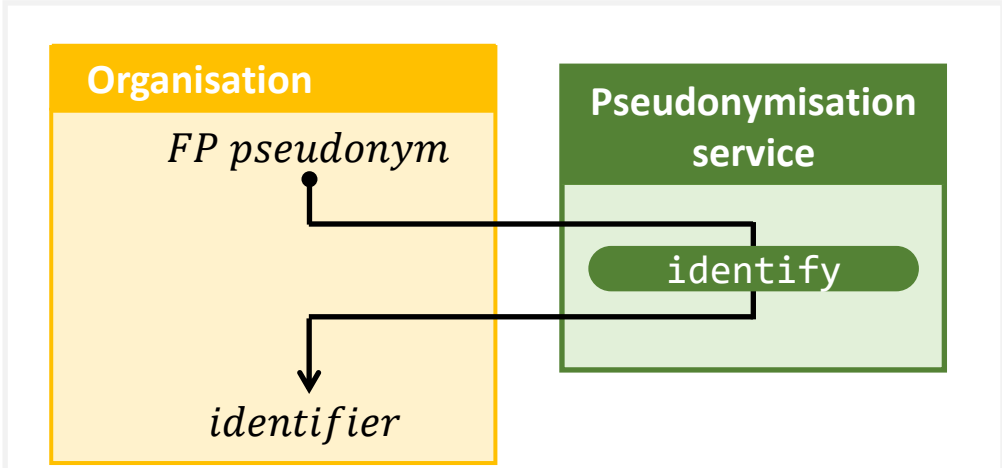
✓ M

Identifiers

✓

✓ Y MB





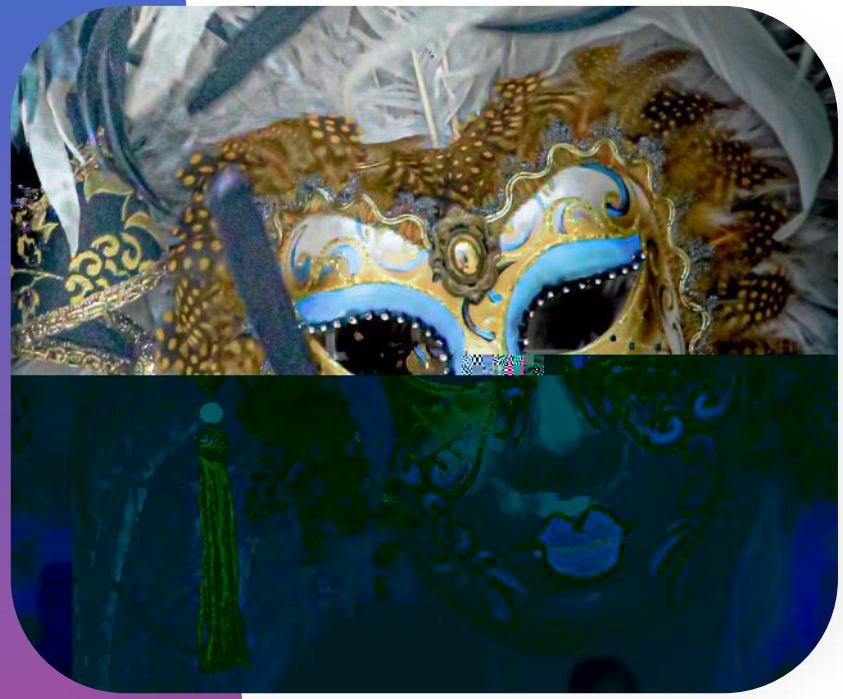
POST Request


```
{
    "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",
        [redacted]
    ]
}
```


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"
        },
```

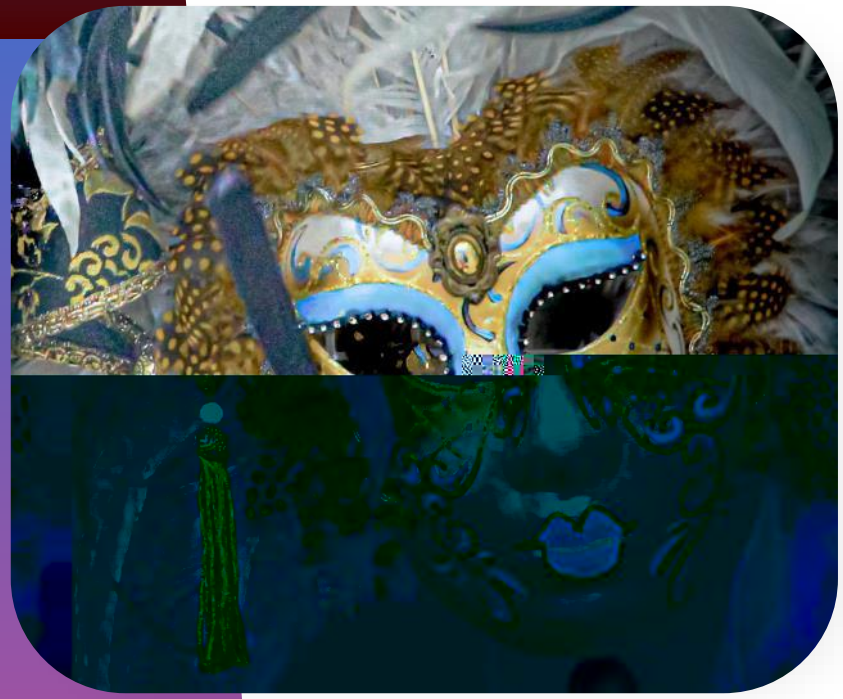
Format-Preserving Pseudonymisation



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

Smals
ICT for society

Format-Preserving Pseudonymisation



Building block

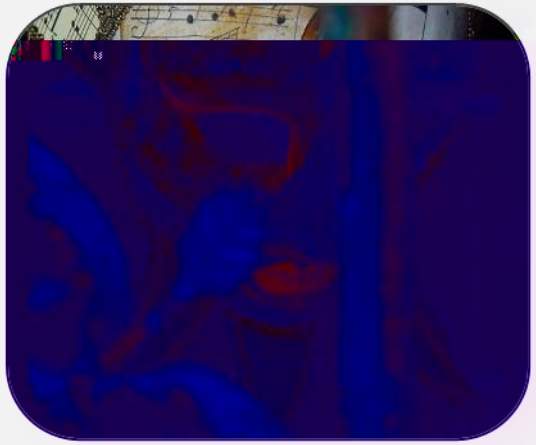
As a Service

Status

Smals
ICT for society

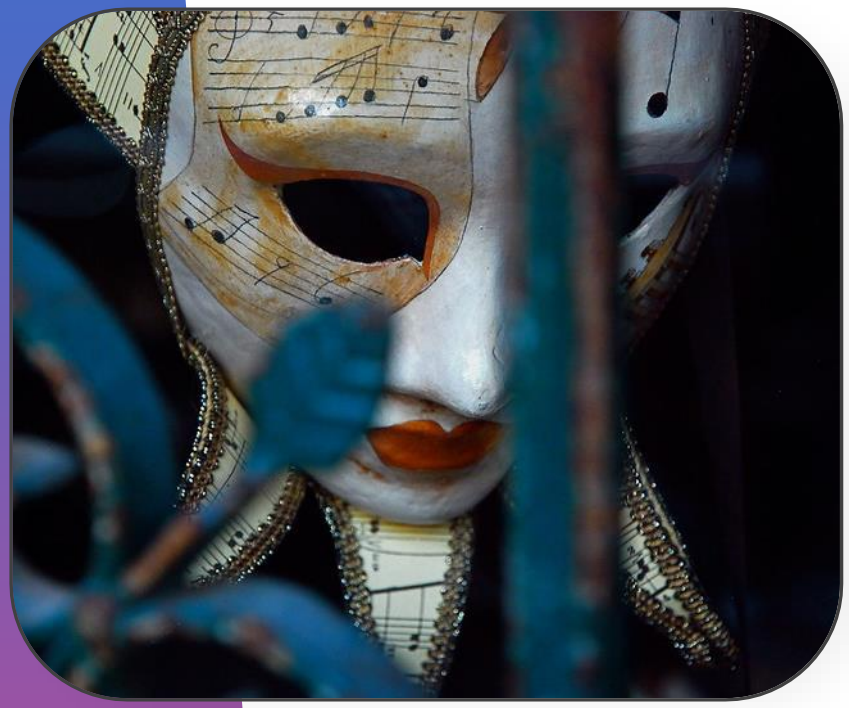
Innovation @ Smals Research

Smart Pseudonymisation



ICT for society

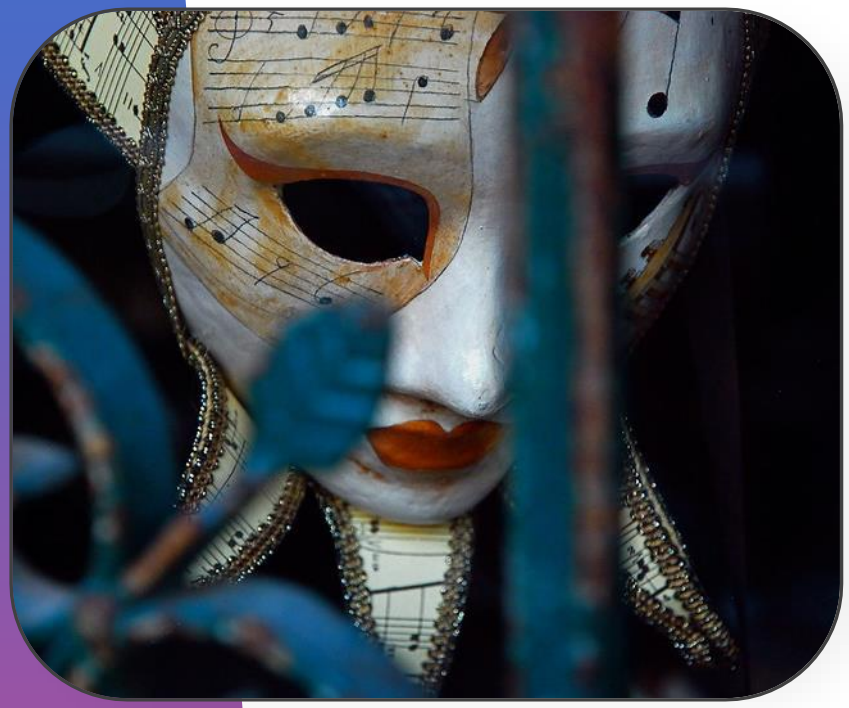
eHealth Blind Pseudonymisation



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

Smals
ICT for society

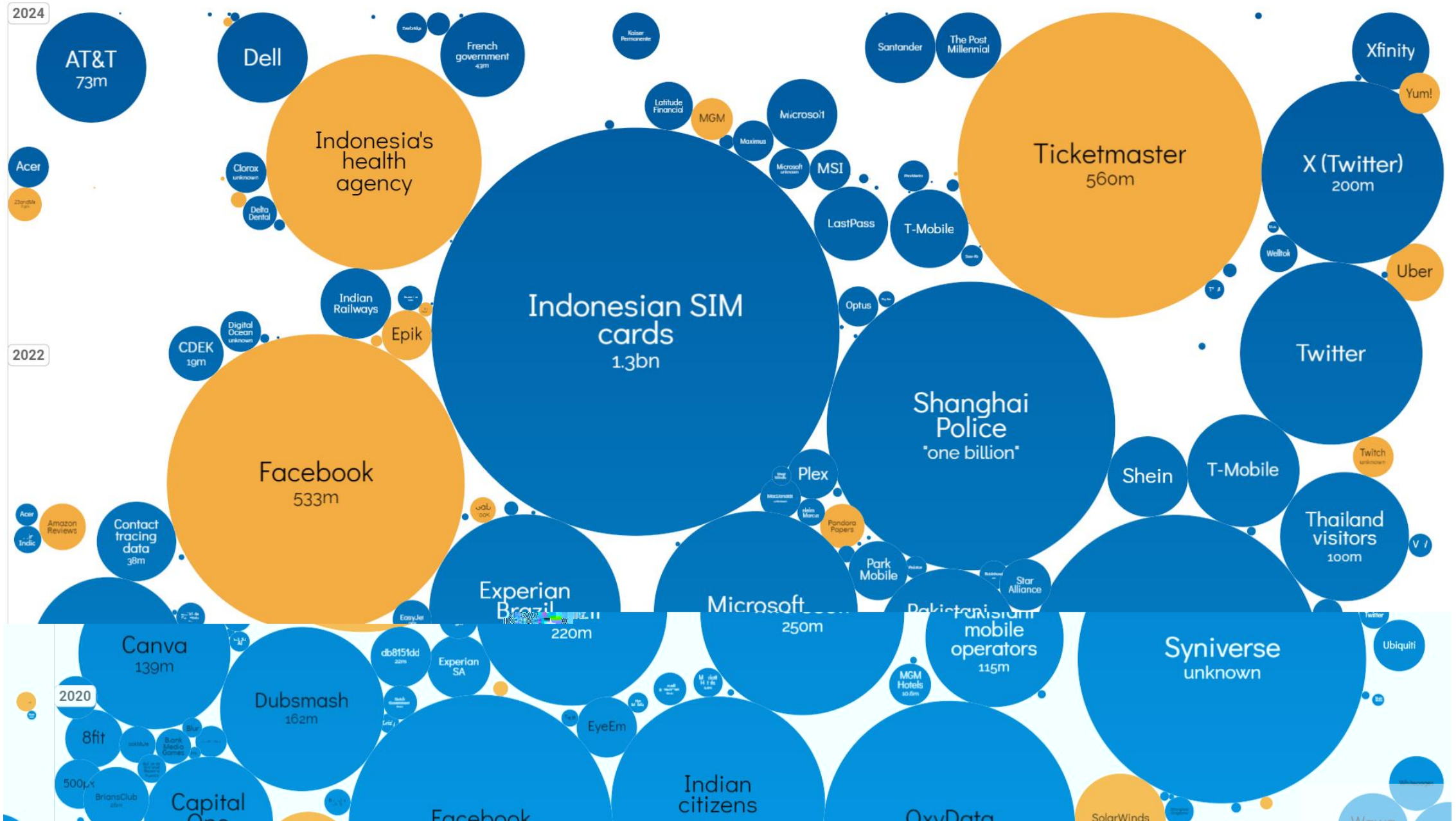
eHealth Blind Pseudonymisation



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

Smals
ICT for society

World's Biggest Data Breaches & Hacks



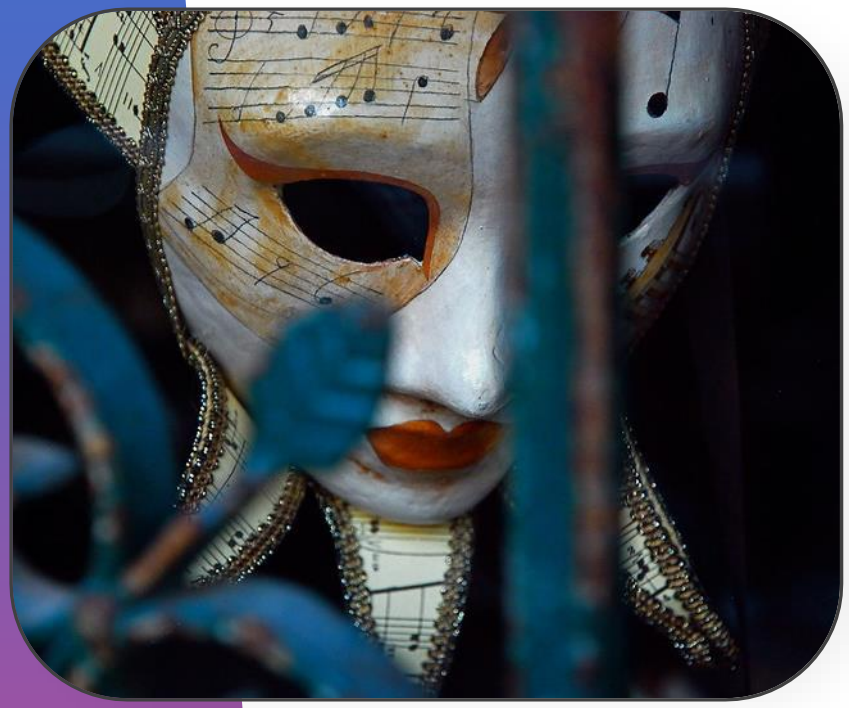
Design principles

Privacy by design

Separation of duties

Simplicity

eHealth Blind Pseudonymisation



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

Smals
ICT for society

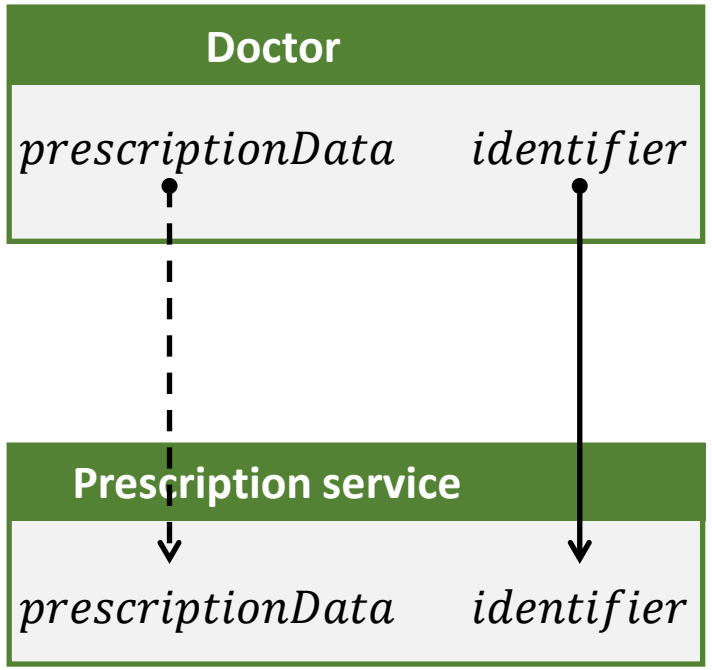
Referral prescription = Verwijsvoorschrift / Prescription de renvoi

What?

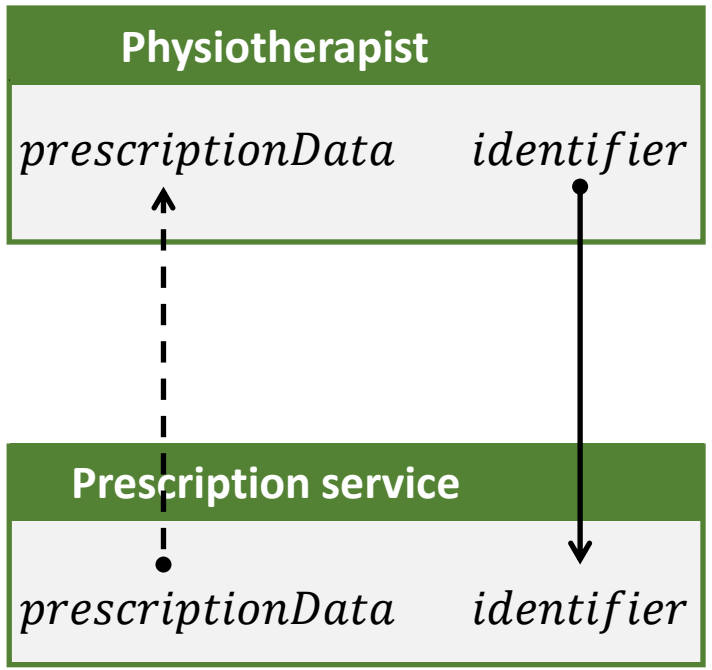
Requirements

- Pseudonymisation
- Partial encryption

Scenario 1



Scenario 2



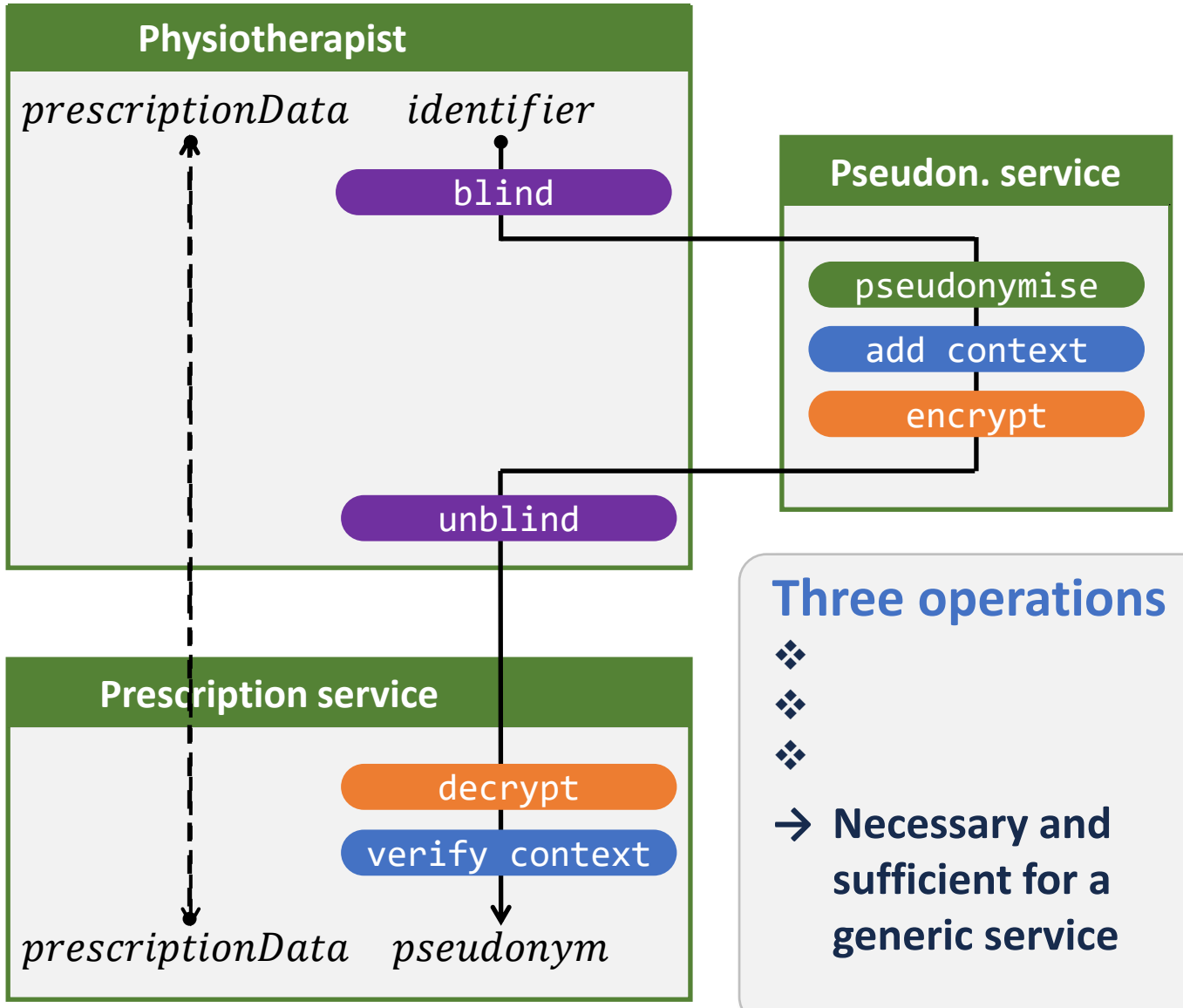
Smals
ICT for society

Blind Pseudo Service

Pseudonymise

Structure blinded identifier, blinded pseudonym and final pseudonym

- ✓ **Each party only sees only what it needs to see**
  - ❖
  - ❖
  - ❖
  - → **Maximizes security & privacy**
- ✓ **Direct communication**
  - ❖
  - ❖
- ✓ **Low-intrusive client-side**
  - ❖
  - ❖



Three operations

- ❖
- ❖
- ❖

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

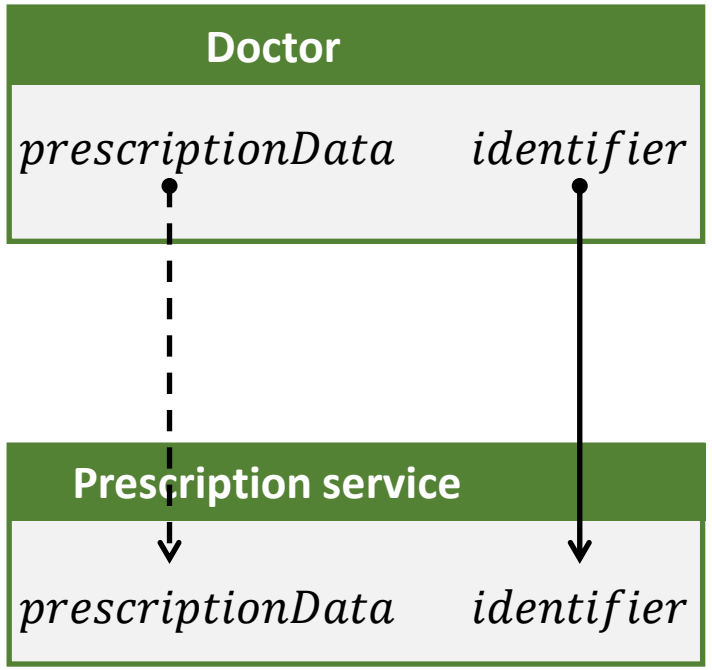
Referral prescription = Verwijsvoorschrift / Prescription de renvoi

What?

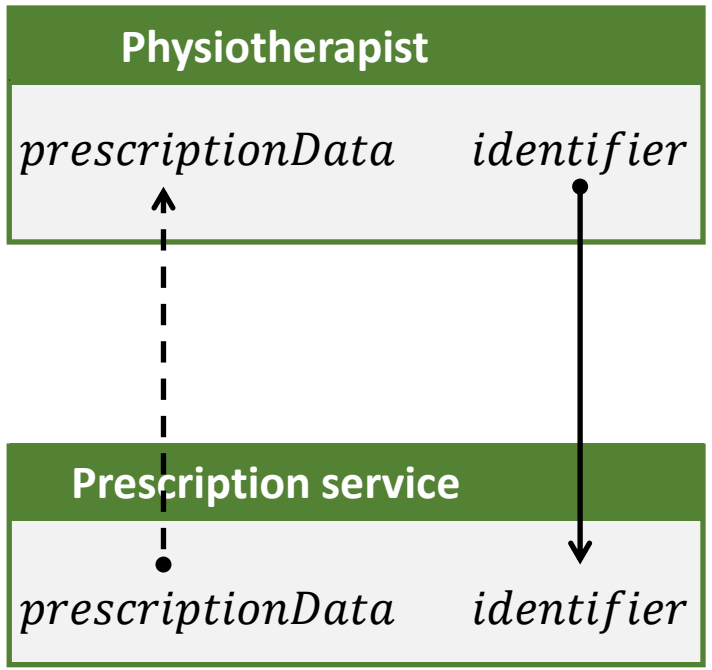
Requirements

- Pseudonymisation
- Partial encryption

Scenario 1



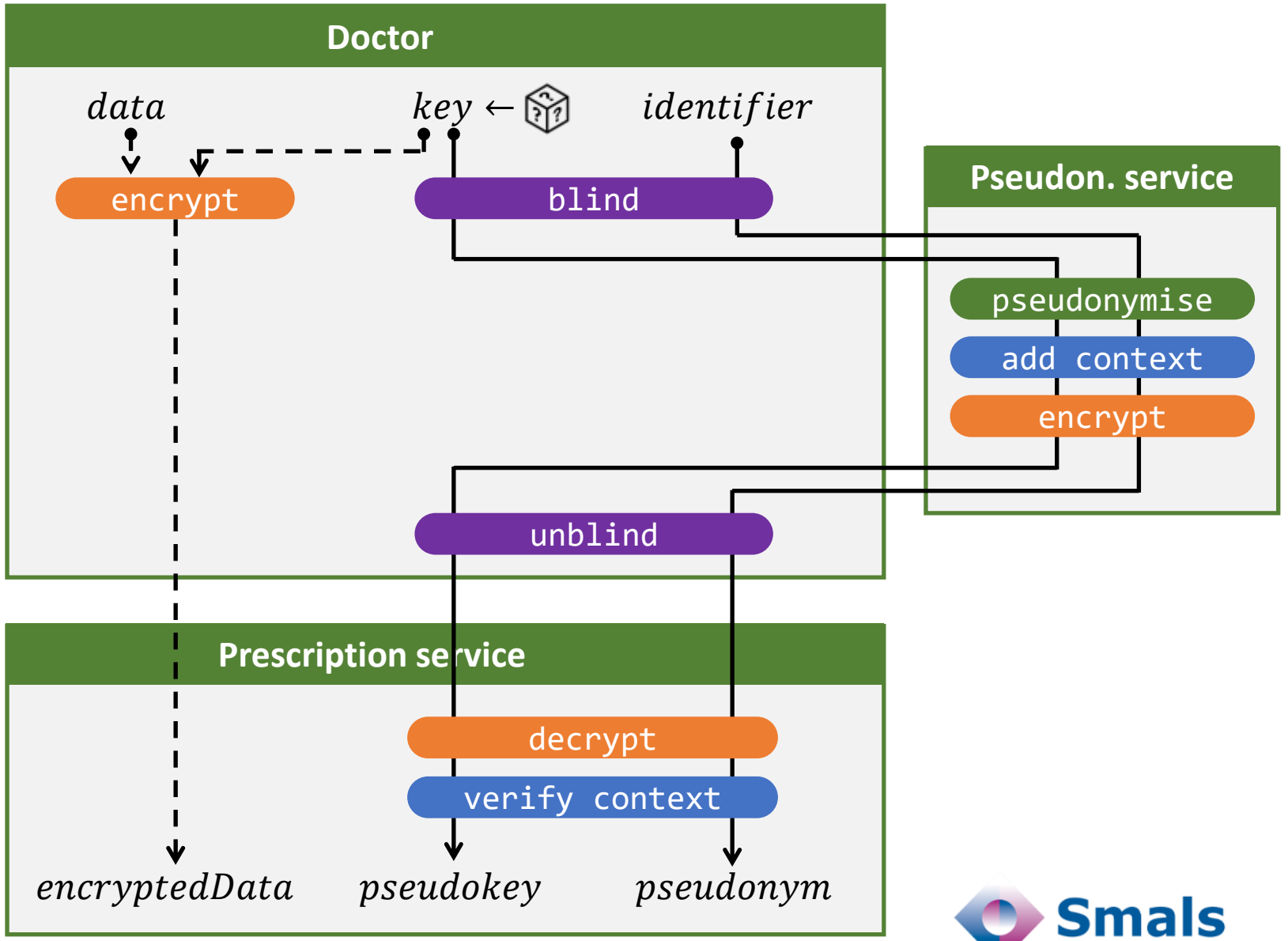
Scenario 2



Smals
ICT for society

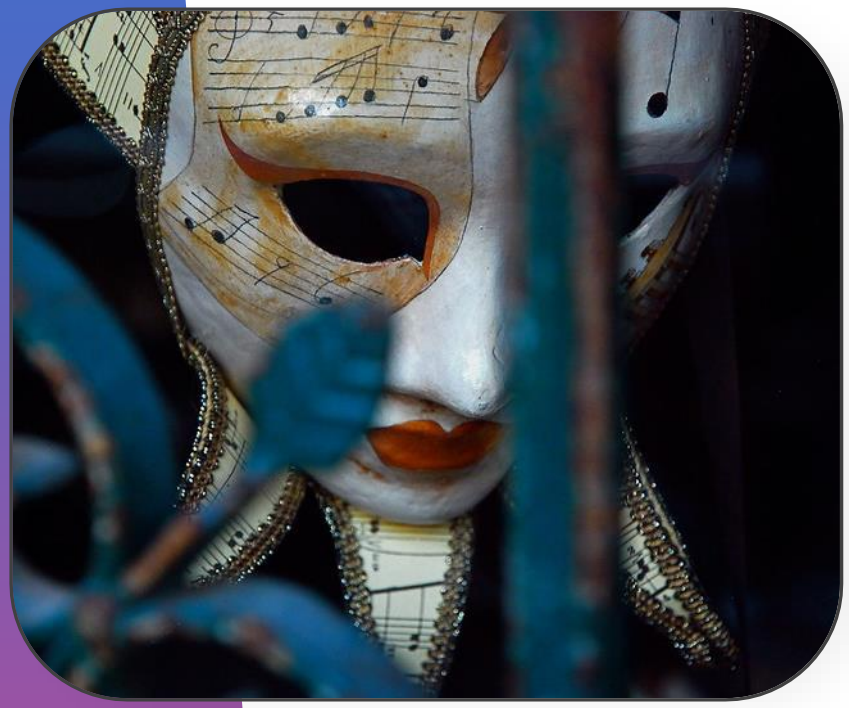
Blind Pseudonymisation Service

Encrypt



Smals
ICT for society

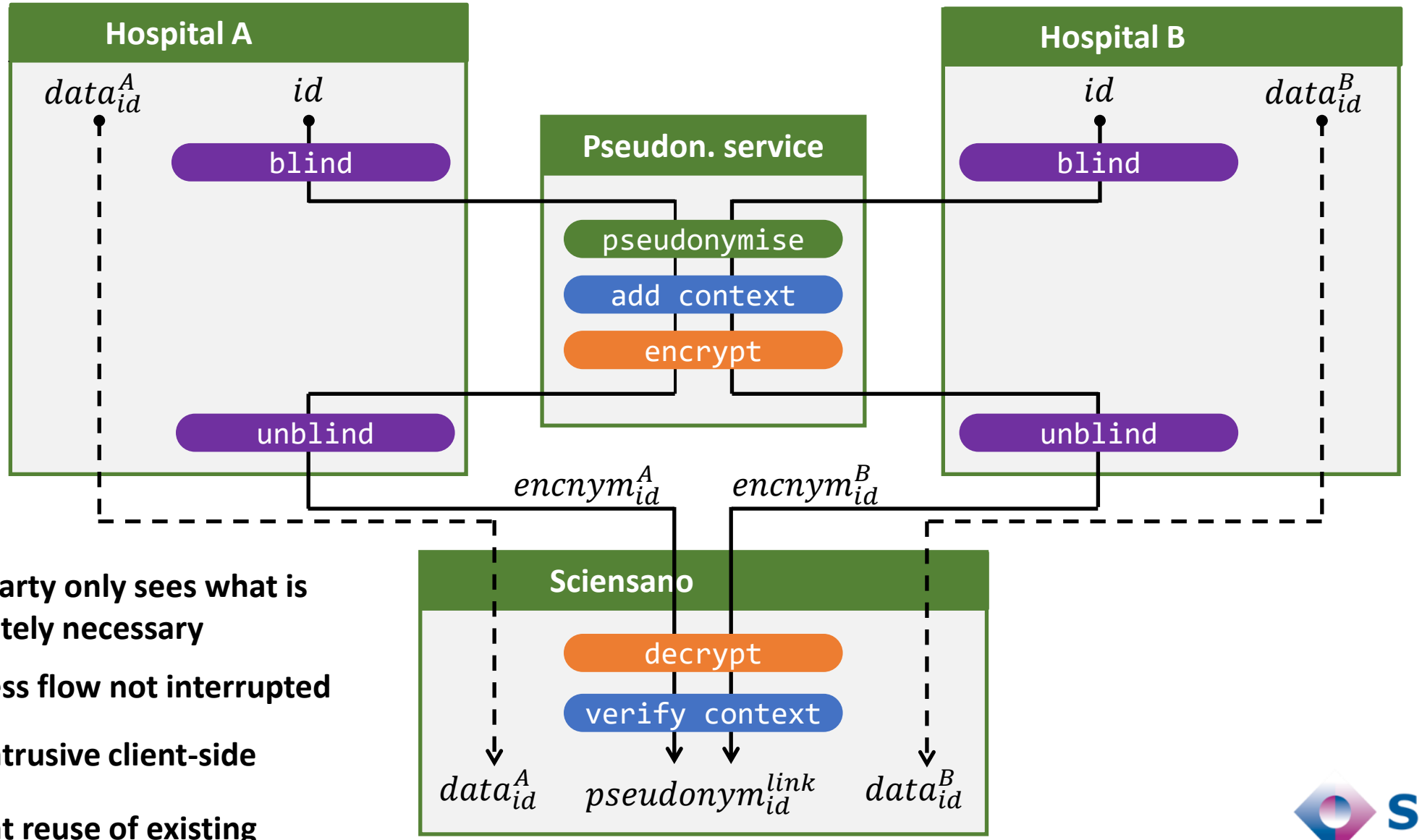
eHealth Blind Pseudonymisation



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

Smals
ICT for society

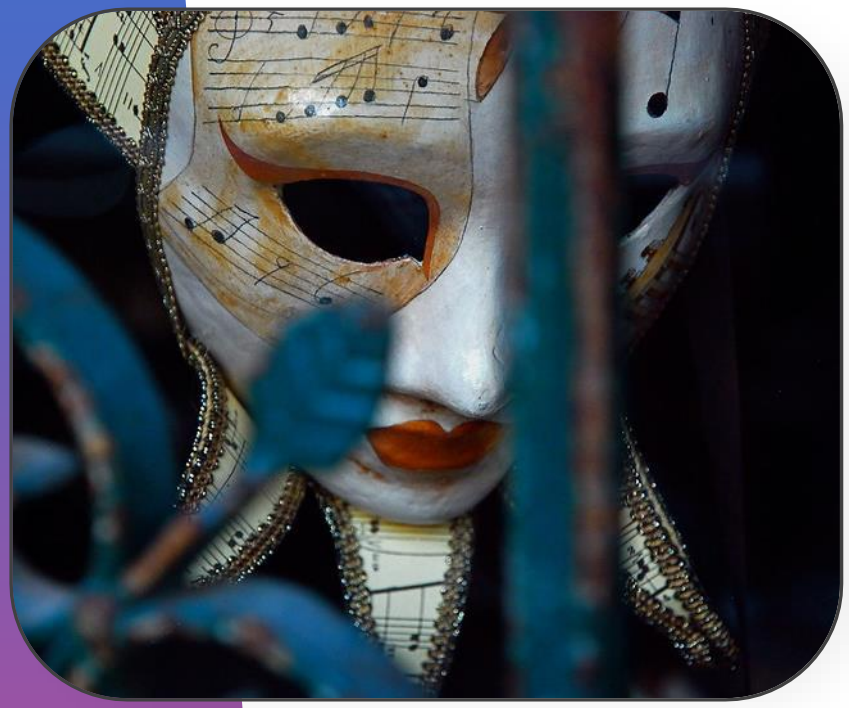
Join & pseudonymise data for research



- ✓ **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

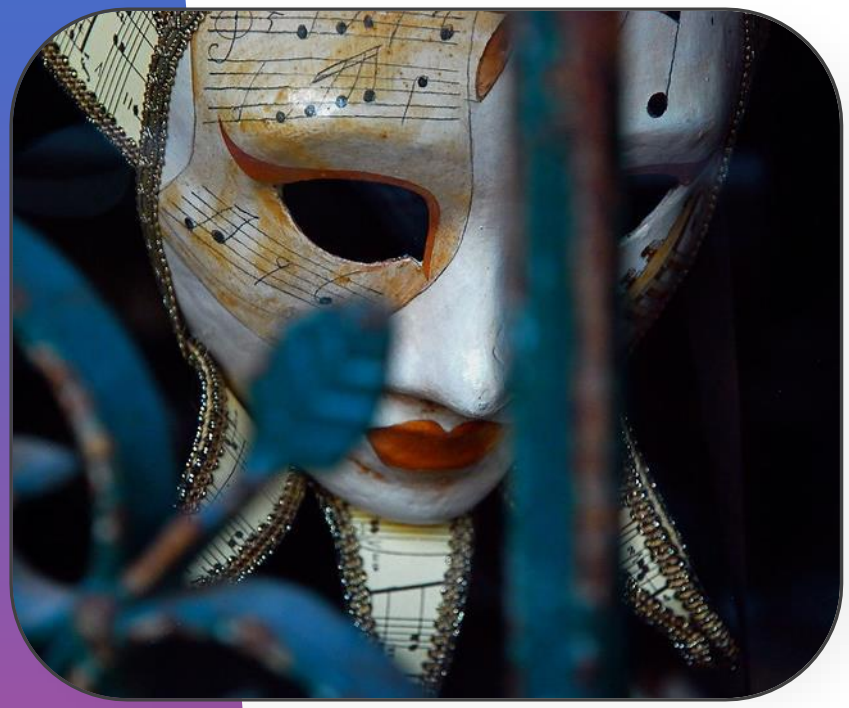
eHealth Blind Pseudonymisation



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

Smals
ICT for society

eHealth Blind Pseudonymisation



Privacy by design

- ✓
- ✓
- ✓

Separation of duties

Simplicity

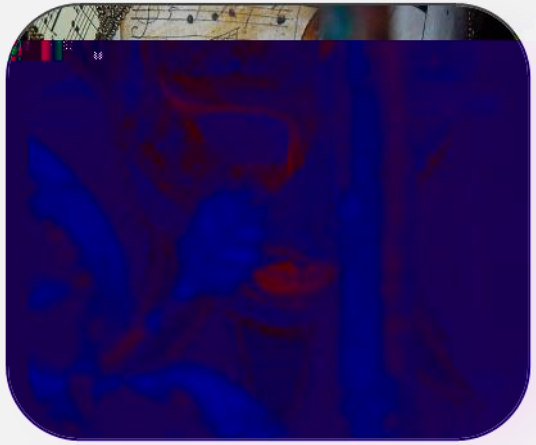
- ✓
- ✓

Live with uptake

Smals
ICT for society

Innovation @ Smals Research

Smart Pseudonymisation



ICT for society

Oblivious Join

- Problem statement
- Concept
- In practice
- Conclusion

Smals
ICT for society

Oblivious Join

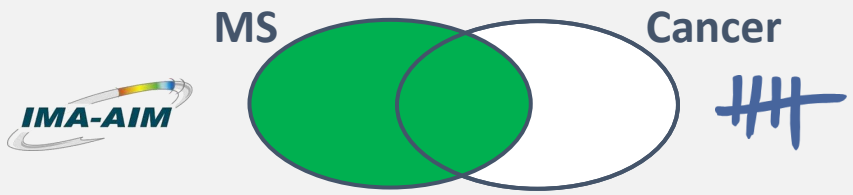
- **Problem statement**
- Concept
- In practice
- Conclusion

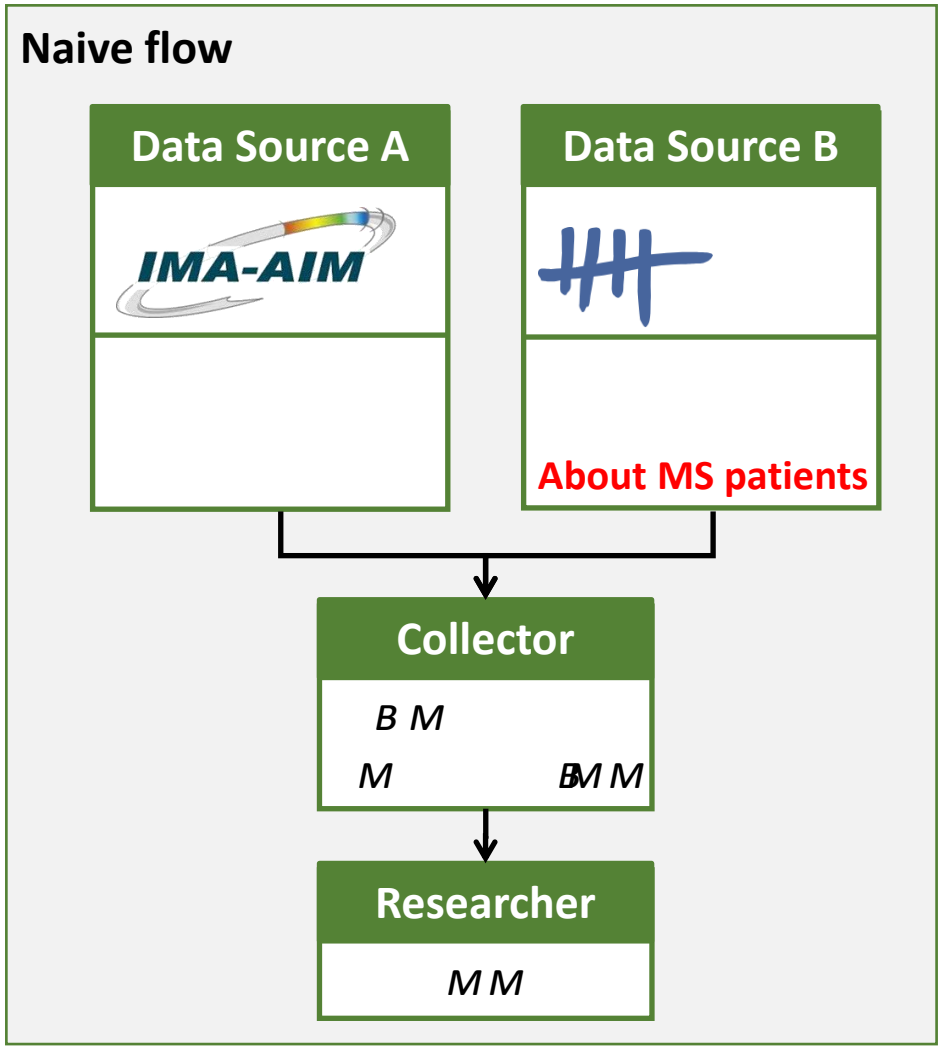
Smals
ICT for society

Concrete case

Research question

Involved citizens





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

Smals
ICT for society

Current practice

Observations

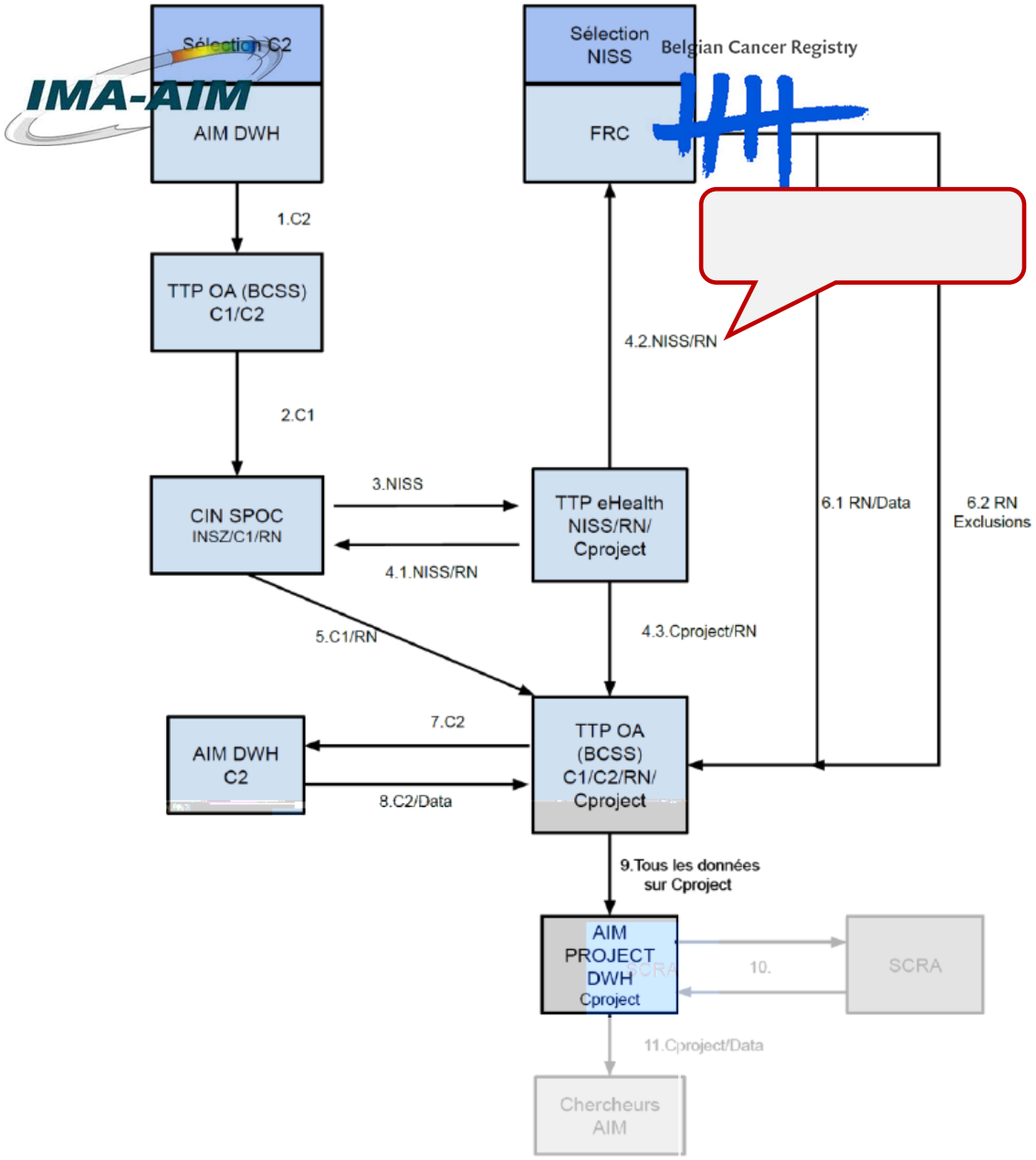
- ✗
- ✗
- ✗
- ✗
- ✗
- ✗

Feedback

MM M M W M

B M Y B M M

Other countries



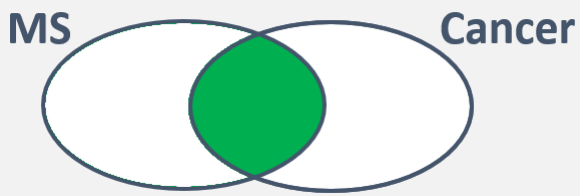
Challenge

Constraint

Requirements

- Privacy-friendly
- Uniform
- No data aggregation
- Easy to use

Focus: set intersection



Extensible from there

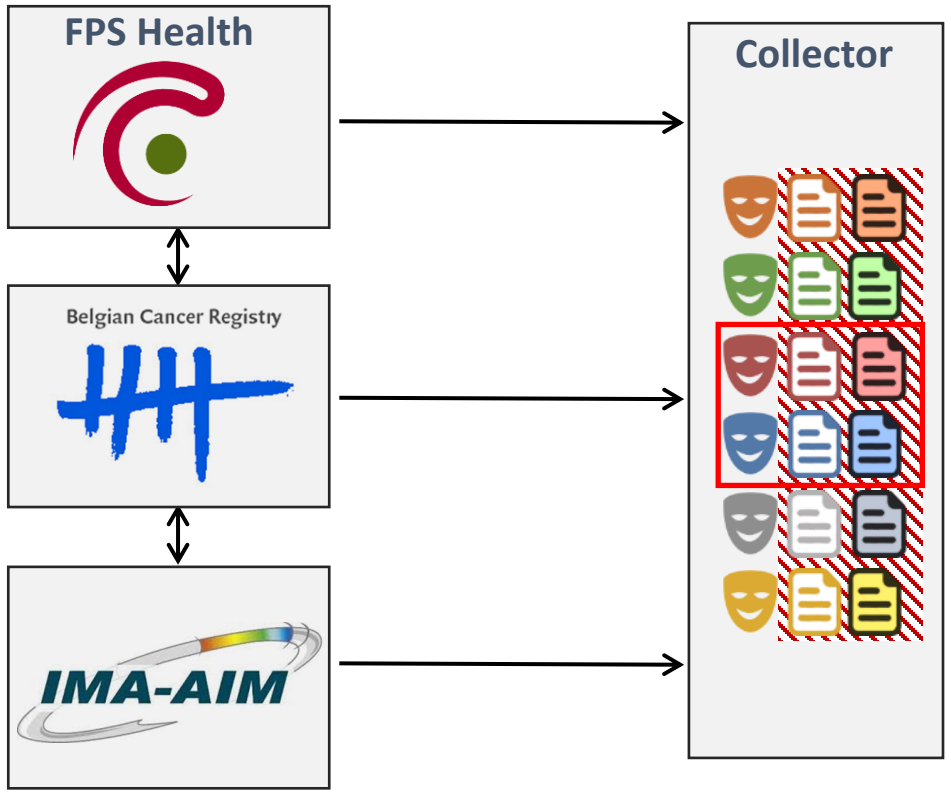
Smals
ICT for society

Oblivious Join

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

Smals
ICT for society

Concept



Properties

- ✓
- ✓
- ✓
- ✓

3 steps protocol

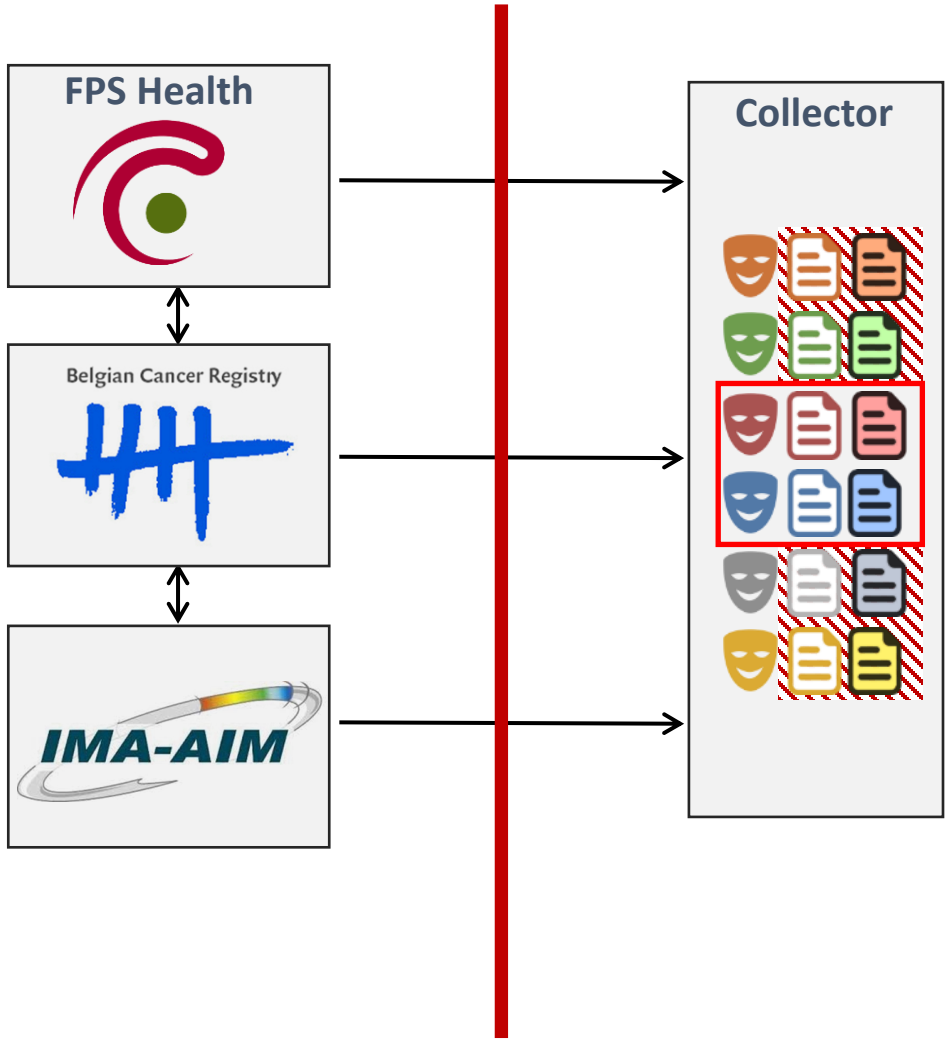
Data sources

- ❖
- ❖

Collector

- ❖
- ❖
- ❖

Concept



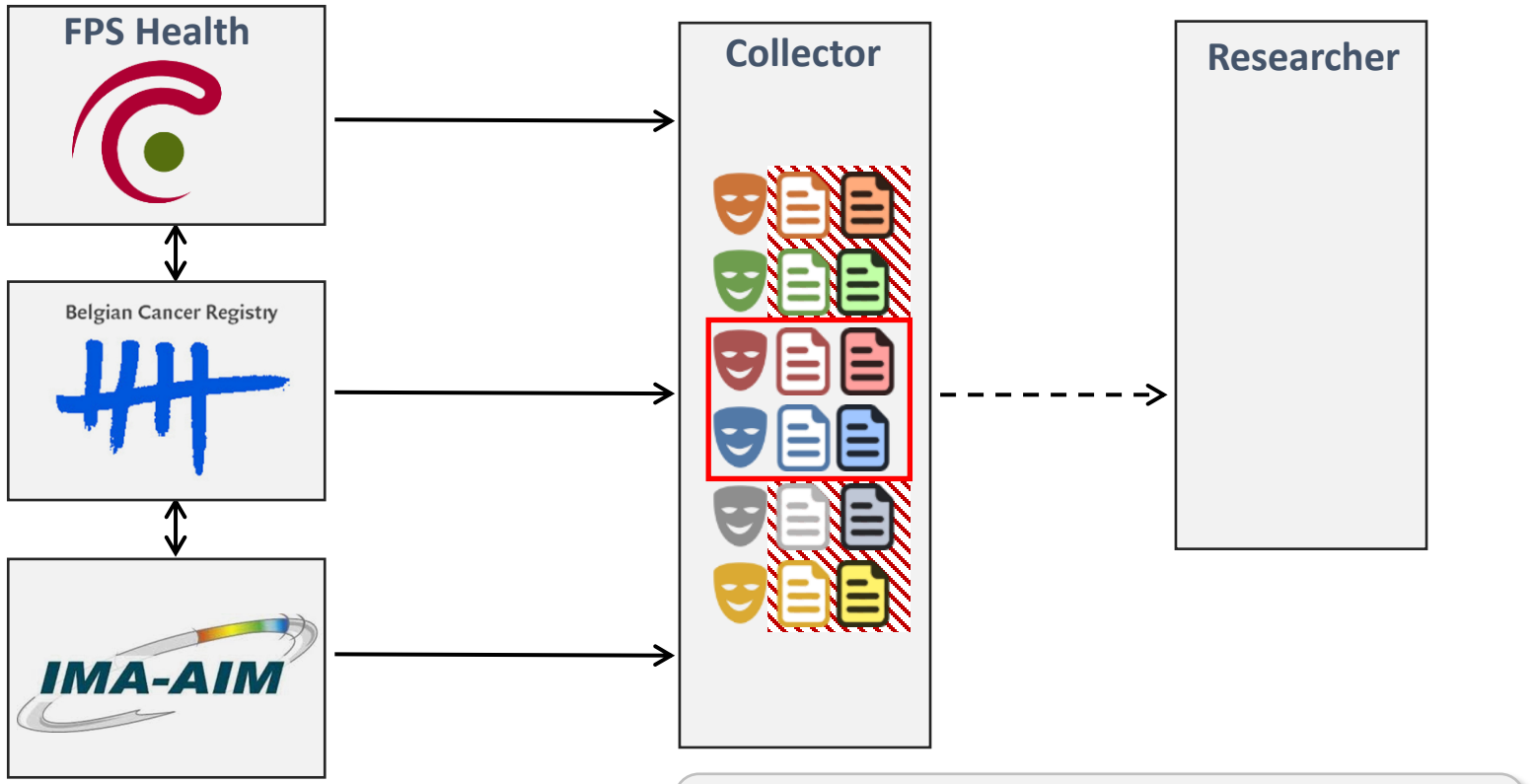
No collusion between data source and collector

Properties

- ✓
- ✓
- ✓
- ✓

3 steps

Concept



Collector

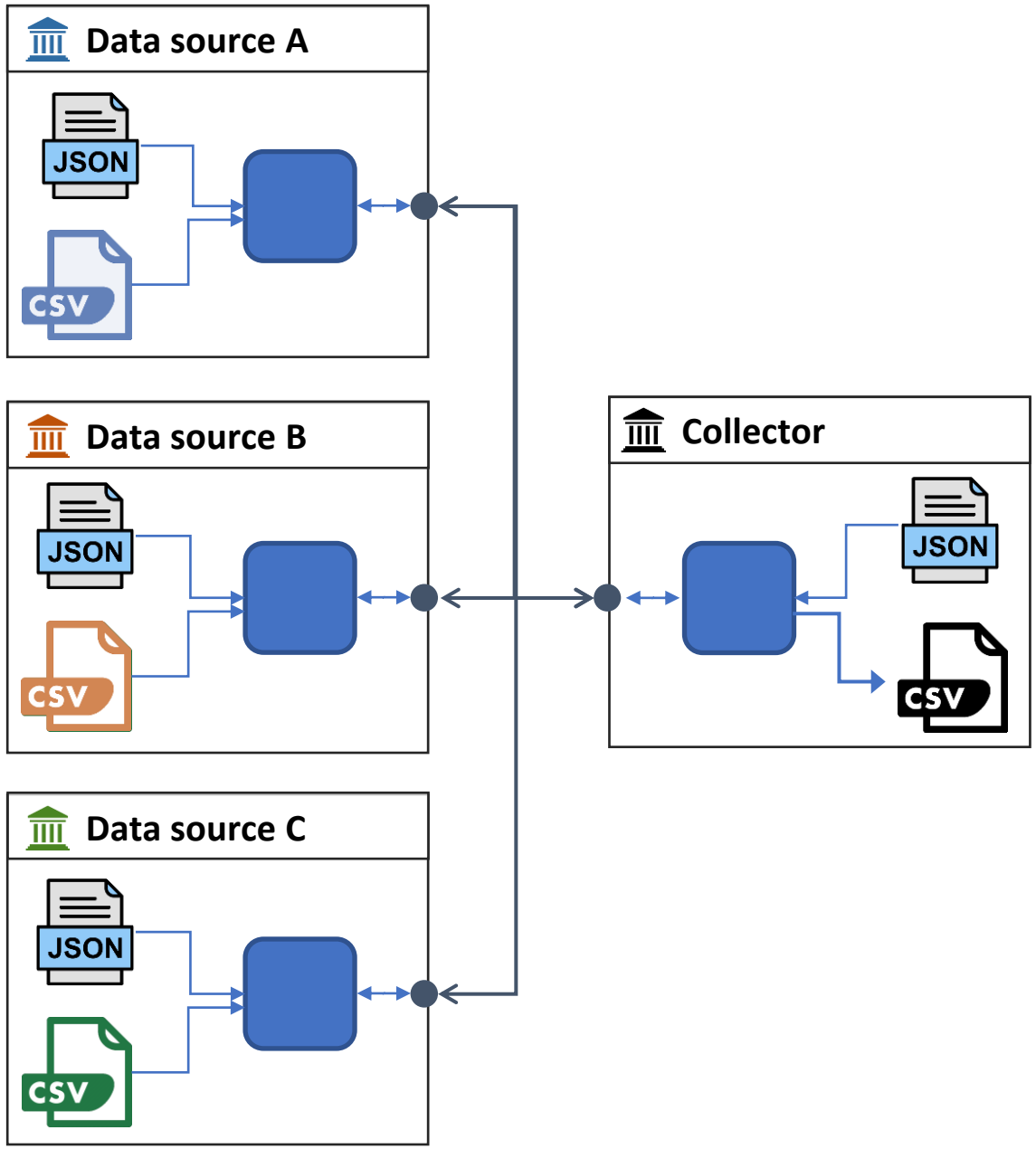
Independent and semi-trusted

Oblivious Join

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

Smals
ICT for society

In practice



Client

Project description

Input files

Output file

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

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

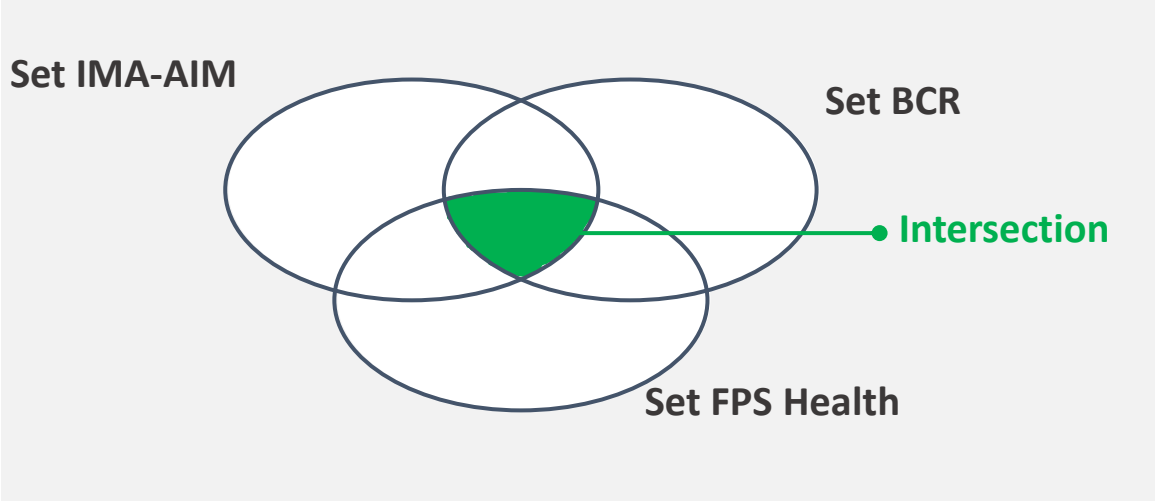
500 000 records

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



Smals
ICT for society

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

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

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

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?

-
-
-

Smals
ICT for society

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

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

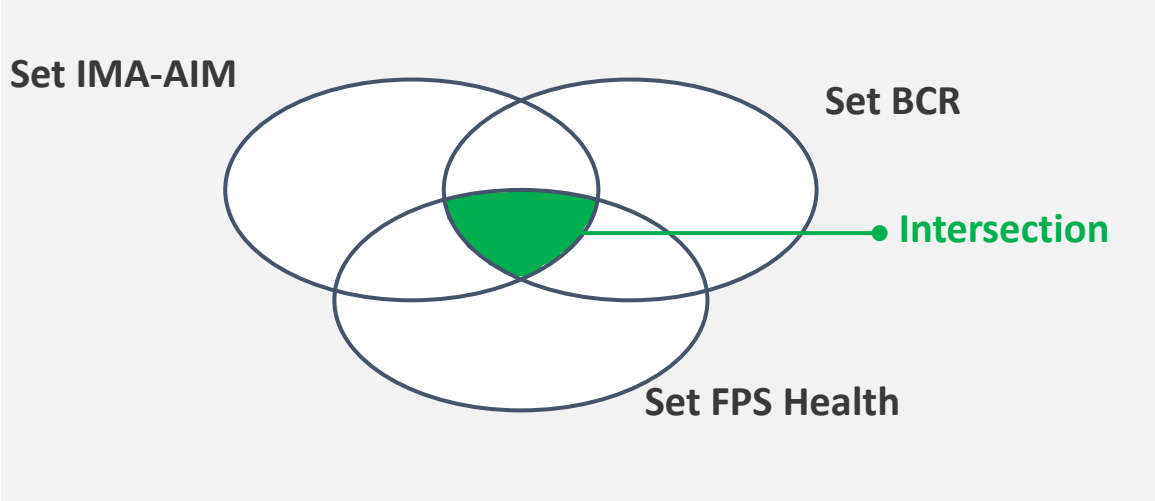
500 000 records

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



Performance test

Parameters

Infrastructure

Results

< 2 min calculations

Oblivious Join

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

Smals
ICT for society

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

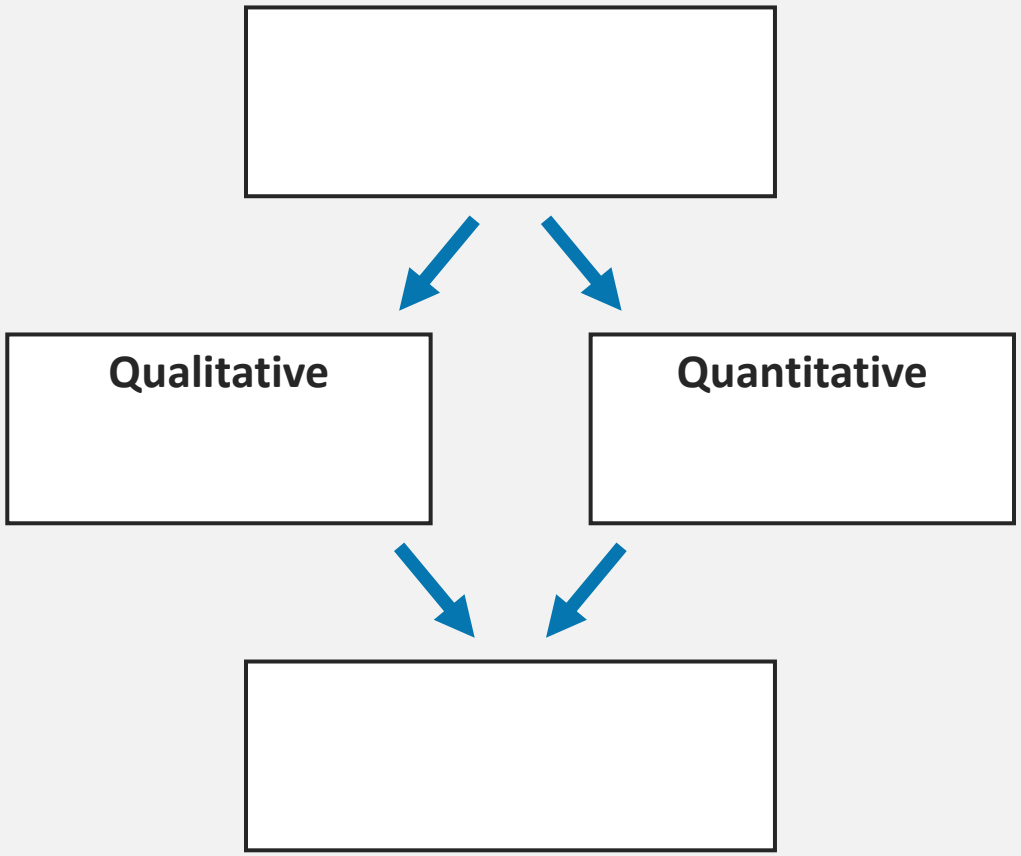
Smals ICT for society

Evaluation

Advantages

Challenges

Opportunities



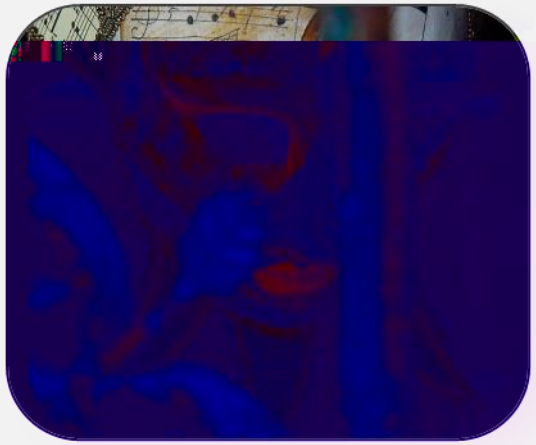
Smals
ICT for society

Wrapping up

Smals
ICT for society

Innovation @ Smals Research

Smart Pseudonymisation



Smart pseudonymisation can play a crucial role to protect personal data

Smals
ICT for society

Thanks for your attention

kristof.verslype@smals.be
+32(0)2 7875376
linkedin.com/in/verslype

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

DEVOXX™

Smals
ICT for society

Further reading
www.smalsresearch.be

www.smalsresearch.be/tag/pseudonymisation/

Images

Judy Dean

Pixabay

Daniel Bruce

estorde

Aris Gionis

Oscar Gende Villar

Smals
ICT for society