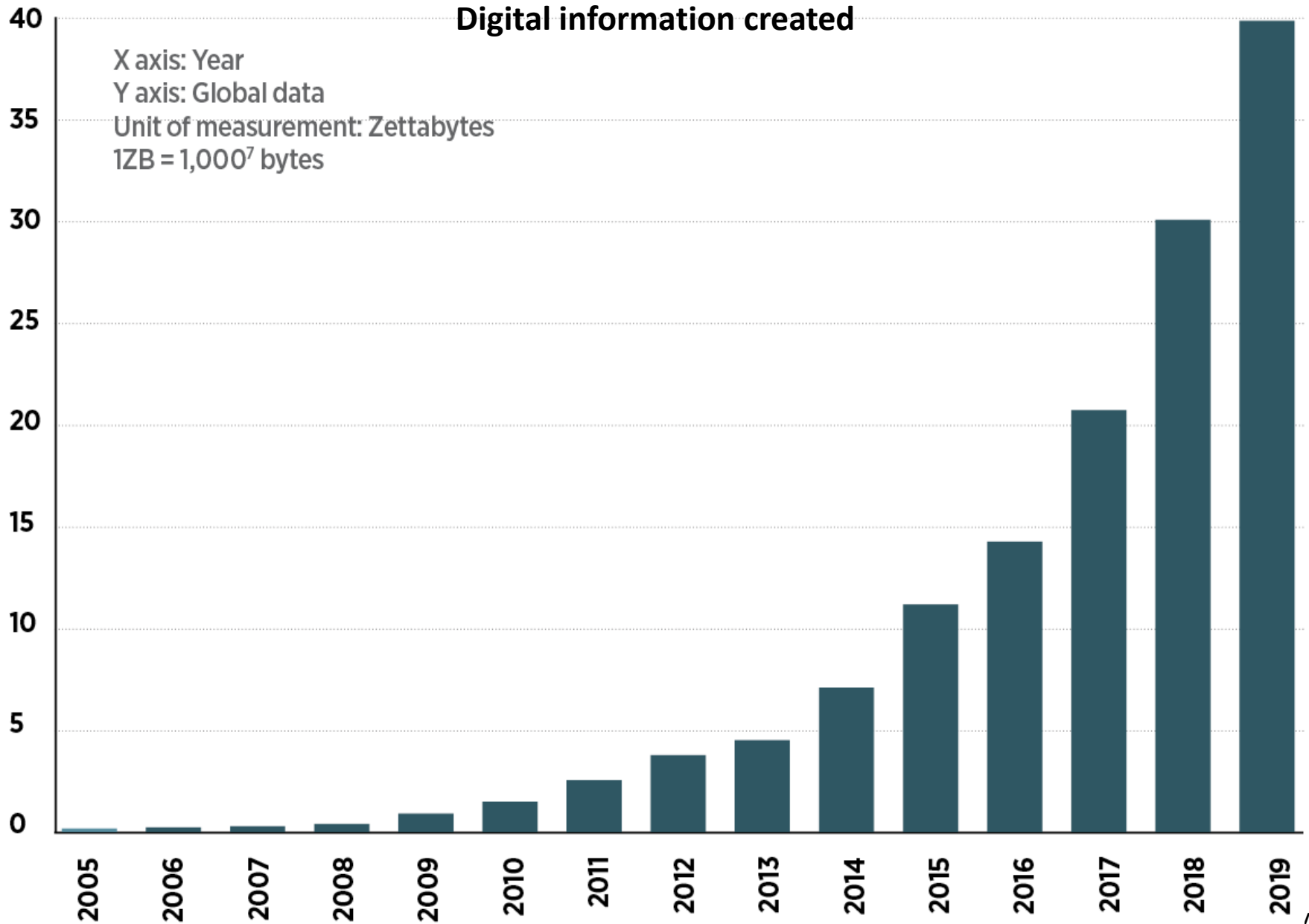


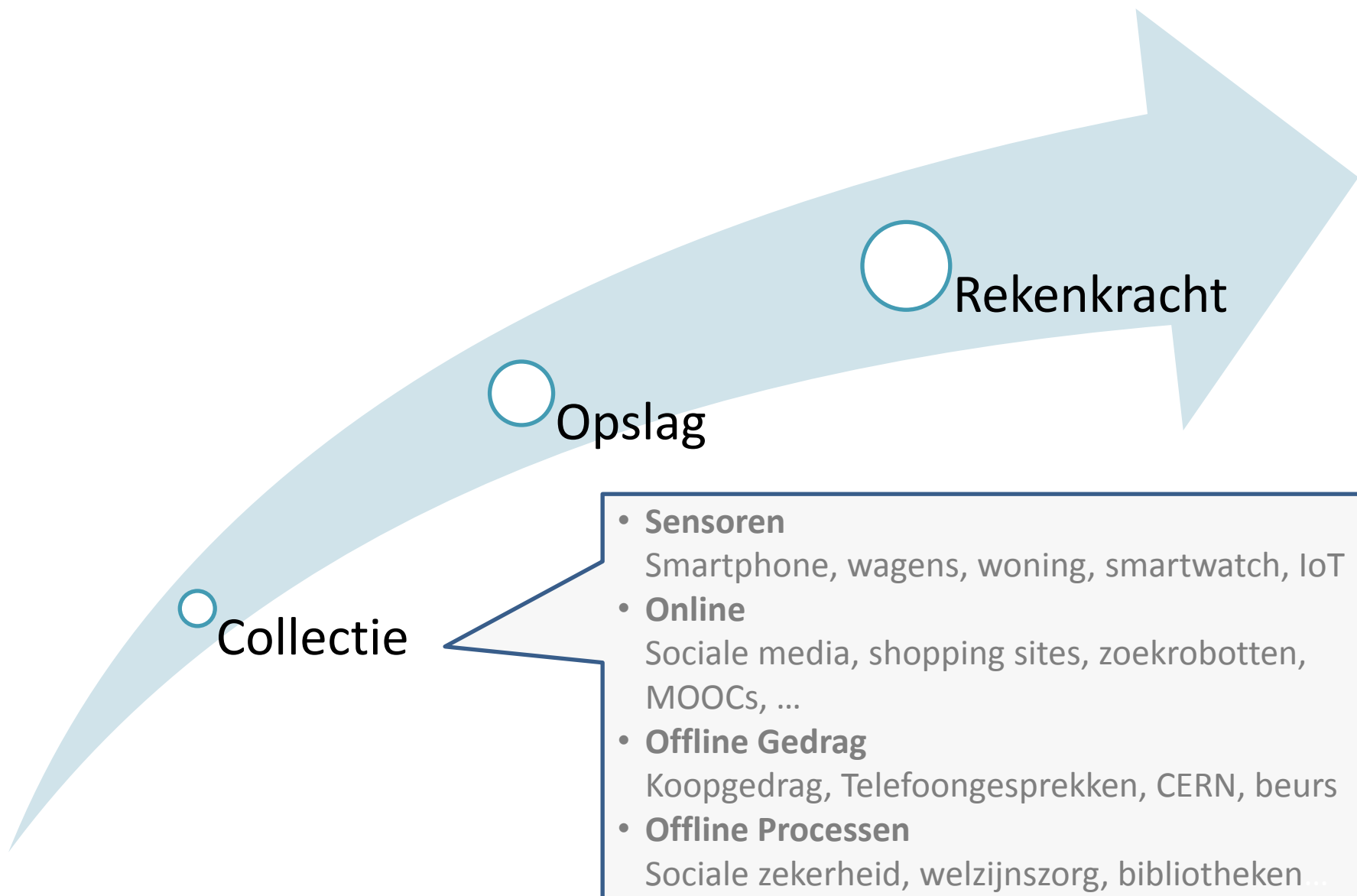
# Privacy vs. Analytics



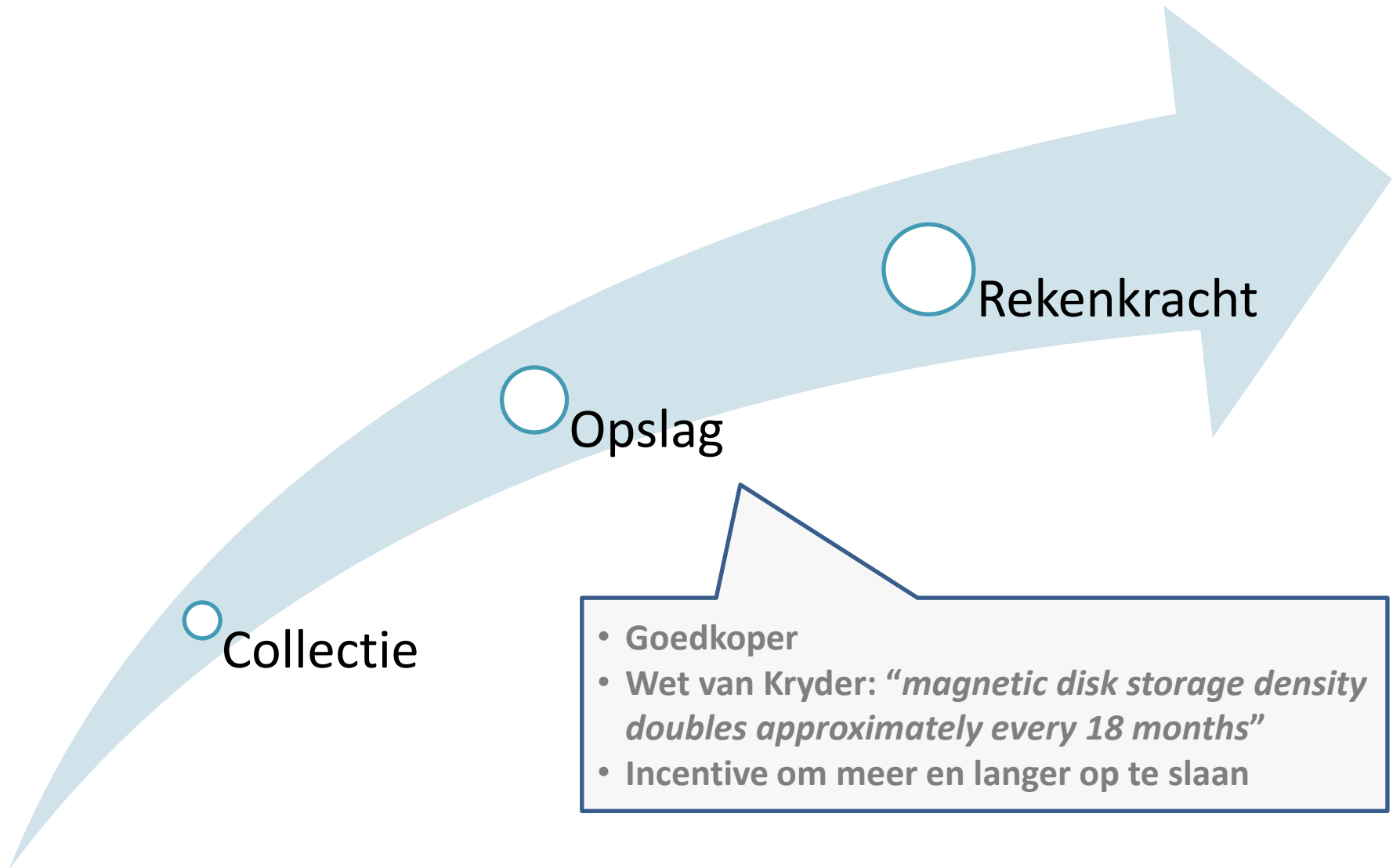
# Data Explosie



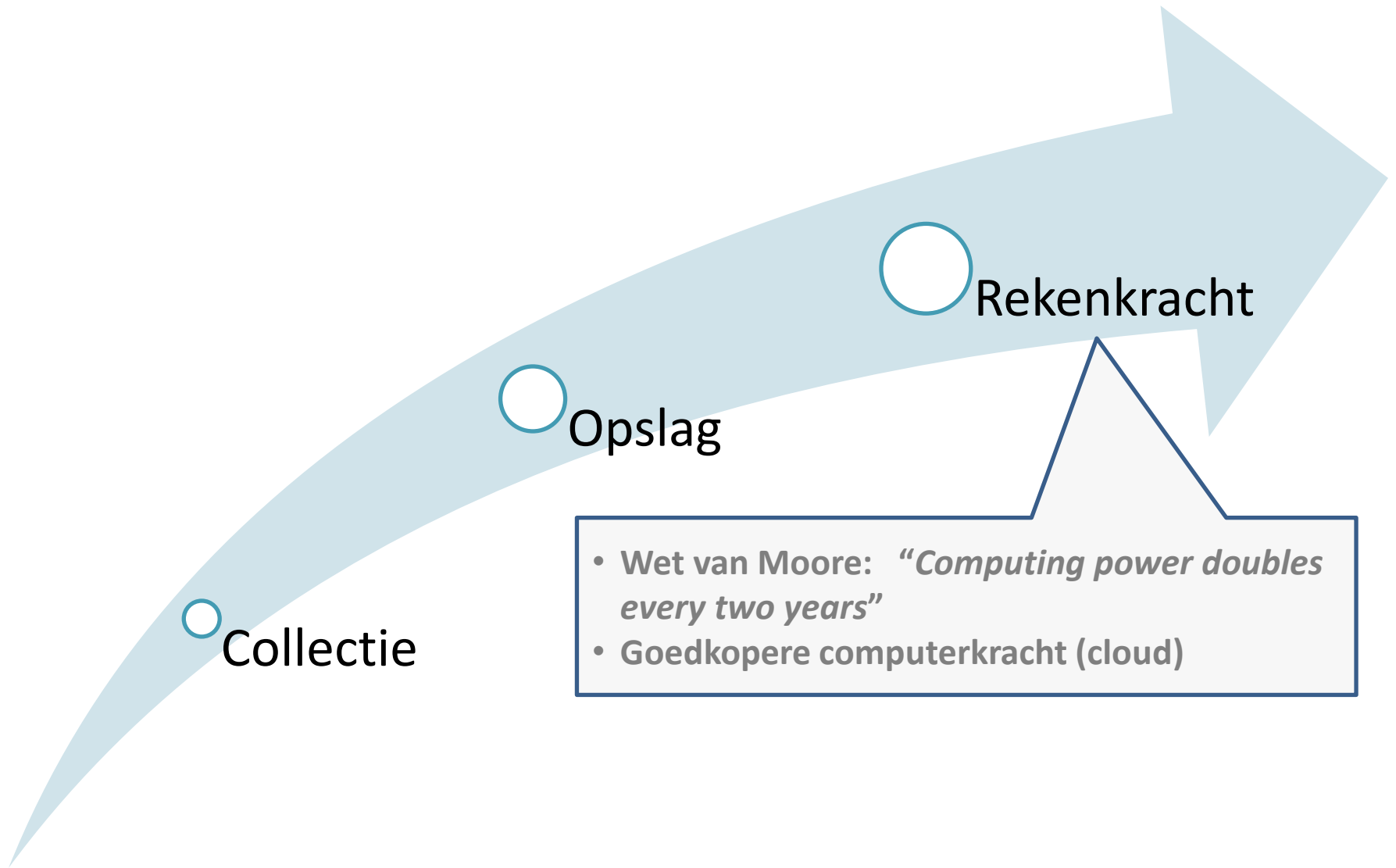
# Data Explosie



# Data Explosie



# Data Explosie

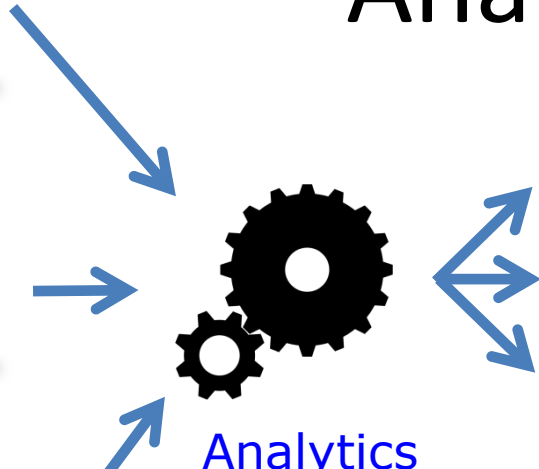


- Wet van Moore: *“Computing power doubles every two years”*
- Goedkopere computerkracht (cloud)

# Analytics



Data



Analytics

Complexe queries,  
Statistiek,  
Modelleren,  
Machine learning, ...

Descriptive analytics

Predictive analytics

Prescriptive analytics

**Enorm potentieel**

**Fraudedetectie**

**Wetenschappelijk  
onderzoek**

**Medische  
analyses**

**Beleids-  
ondersteuning**

**Rechten / premies  
burger**

...

In geval van persoonsgegevens is er...

# De privacywet

(aka Wet tot bescherming van de persoonlijke levensfeer  
ten opzichte van de verwerking van persoonsgegevens (8/12/1992 )



# Privacywet

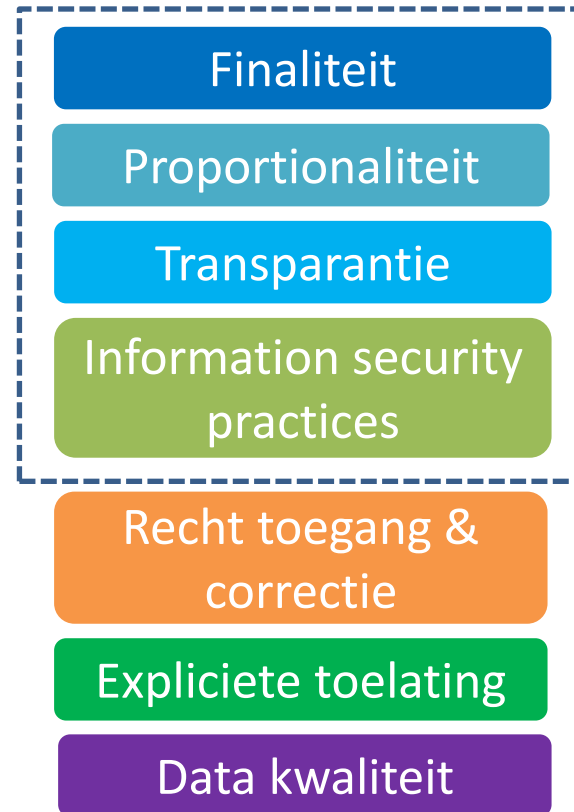
**Anonieme gegevens**

Onidentificeerbaar

**Persoonsgegevens**

Identificeerbaar/  
Geïdentificeerd

Privacywet niet  
van toepassing



**Waarschijnlijkheid op identificatie**

# Privacywet

Anonieme gegevens

Onidentificeerbaar

Persoonsgegevens

Identificeerbaar/  
Geïdentificeerd

## Identificeerbaar:

- Direct: identificatienummer  
vb. INSZ-nummer
- Indirect: a.d.h.v. één of meer specifieke elementen die kenmerkend zijn  
vb. geslacht, postcode en geboortedatum

Recht toegang &

correctie

Expliciete toelating

Finaliteit

Transparantie

Information security,

practices

Data kwaliteit

Waarschijnlijkheid op identificatie

# Kruisen van Persoonsgegevens

## Een fictief voorbeeld

*Een onderzoeksteam wil medische, financiële en demografische persoonsgegevens analyseren van alle burgers die na 1990 geboren zijn, die zelfstandige in bijberoep zijn met een loon van minstens € 50 000.*

*Deze gegevens worden echter beheerd door verschillende overheidsbedrijven en moeten dus gekruist worden.*

Kan technologie steeds persoonsgegevens converteren naar anonieme gegevens die vervolgens geanalyseerd kunnen worden?



Wat is er technisch mogelijk binnen het wettelijke kader?

# AGGENDA

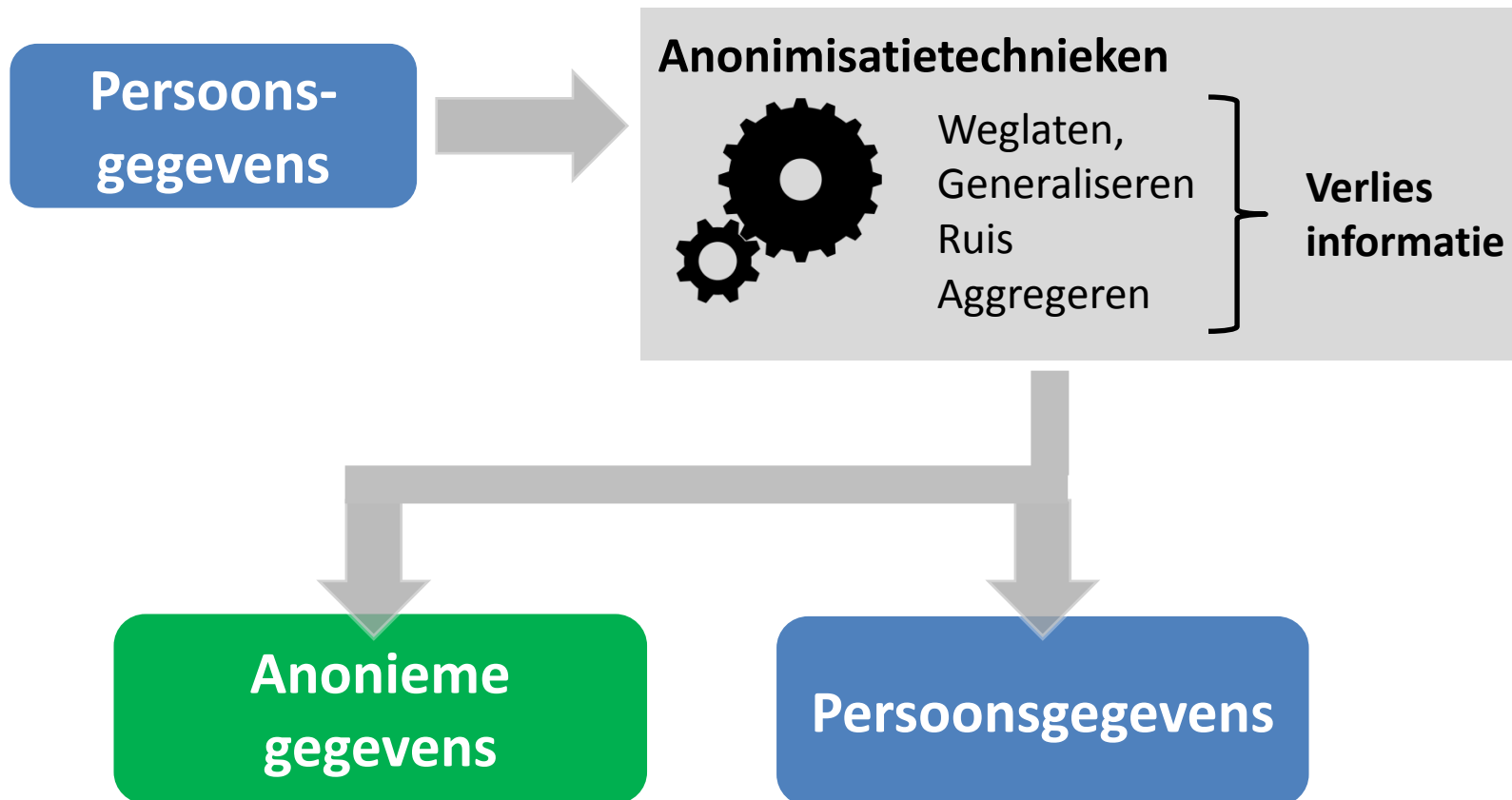
- **Introductie**
- Technieken Anonimisatie**
  - Wat
  - Beperkingen
- Data Archipel**
  - Concept
  - Proof of concept
  - Een beetje crypto
  - Deanonimisatie
  - Sleutelbeheer
  - Transparantie
  - Small cells
- Conclusies**



# Anonimisatie- technieken

# Anonimisatietechnieken

*Transformeren van persoonsgegevens  
zodat identificatie moeilijker wordt*



Toepassen anonimisatietechnieken  $\nRightarrow$  juridisch anonieme gegevens

# Verwijderen IDs

Naam	RRN	DoB	Sex	ZIP	Problem
Kristof Verslype	84052010955	20/05/1984	M	3211	Antisocial personality disorder
Jan Vandesmals	76011320675	13/01/1976	M	9300	Schizofrenie
Paola Ruffo di Calabria	37091100247	11/09/1937	F	1060	Perfect gezond
Dick Tatuur	50111221385	12/11/1950	M	4000	Megalomanie



Naam	RRN	DoB	Sex	ZIP	Problem
████████████████████	████████████████████	20/05/1984	M	3211	Antisocial personality disorder
████████████████████	████████████████████	13/01/1976	M	9300	Schizofrenie
████████████████████	████████████████████	11/09/1937	F	1060	Perfect gezond
████████████████████	████████████████████	12/11/1950	M	4000	Megalomanie

# Coderen IDs

RRN	Code
84052010955	X38LS45
76011320675	X38DI56
37091100247	X38XD12
50111221385	X38MP68

Naam	RRN	DoB	Sex	Z	
Kristof Verslype	84052010955	20/05/1984	M	3211	Antisocial personality disorder
Jan Vandesmals	76011320675	13/01/1976	M	9300	Schizofrenie
Paola Ruffo di Calabria	37091100247	11/09/1937	F	1060	Perfect gezond
Dick Tatuur	50111221385	12/11/1950	M	4000	Megalomanie



Naam	RRN	DoB	Sex	ZIP	Problem
	X38LS45	20/05/1984	M	3211	Antisocial personality disorder
	X38DI56	13/01/1976	M	9300	Schizofrenie
	X38XD12	11/09/1937	F	1060	Perfect gezond
	X38MP68	12/11/1950	M	4000	Megalomanie

# Verwijderen Quasi IDs

Naam	RRN	DoB	Sex	ZIP	Problem
X38LS45		20/05/1984	M	3211	Antisocial personality disorder
X38DI56		13/01/1976	M	9300	Schizofrenie
X38XD12		11/09/1937	F	1060	Perfect gezond
X38MP68		12/11/1950	M	4000	Megalomanie



Naam	RRN	DoB	Sex	ZIP	Problem
X38LS45		2 [REDACTED]	M	3211	Antisocial personality disorder
X38DI56		1 [REDACTED]	M	9300	Schizofrenie
X38XD12		1 [REDACTED]	F	1060	Perfect gezond
X38MP68		1 [REDACTED]	M	4000	Megalomanie

# Vervagen

Naam	RRN	DoB	Sex	ZIP	Problem
X38LS45		20/05/1984	M	3211	Antisocial personality disorder
X38DI56		13/01/1976	M	9300	Schizofrenie
X38XD12		11/09/1937	F	1060	Perfect gezond
X38MP68		12/11/1950	M	4000	Megalomanie



Naam	RRN	Age	Sex	Loc	Problem
X38LS45		32	M	VI-B	Antisocial personality disorder
X38DI56		39	M	O-VI	Schizofrenie
X38XD12		77	F	BXL	Perfect gezond
X38MP68		64	M	LUIK	Megalomanie

# Ruis toevoegen

Naam	RRN	DoB	Sex	ZIP	Problem
X38LS45		20/05/1984	M	3211	Antisocial personality disorder
X38DI56		13/01/1976	M	9300	Schizofrenie
X38XD12		11/09/1937	F	1060	Perfect gezond
X38MP68		12/11/1950	M	4000	Megalomanie



Naam	RRN	DoB	Sex	ZIP	Problem
X38LS45		07/07/1984	M	3211	Antisocial personality disorder
X38DI56		28/12/1975	M	9300	Schizofrenie
X38XD12		29/09/1937	F	1060	Perfect gezond
X38MP68		07/12/1950	M	4000	Megalomanie

# Aggregeren

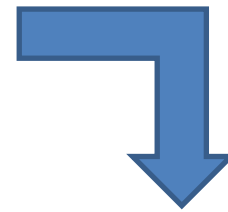
Naam	RRN	DoB	Sex	ZIP	Problem
X38LS45		20/05/1984	M	3211	Antisocial personality disorder
X38DI56		13/01/1976	M	9300	Schizofrenie
X38XD12		11/09/1937	F	1060	Perfect gezond
X38MP68		12/11/1950	M	4000	Megalomanie



DoB	Sex	ZIP	Problem
≥ 1970	...	...	Antisocial personality disorder, Schizofrenie
< 1970	...	...	Megalomanie, Perfect gezond

SSN	Race	BirthDate	Gender	ZIP	Problem
141	021-57-1445	black	9/20/1965	male	02 short of breath
141	021-77-8034	black	2/14/1965	male	02 chest pain
138	107-21-0876	black	10/23/1965	female	02 painful eye
138	021-37-1573	black	8/24/1965	female	02 wheezing
138	021-54-4229	black	11/7/1964	female	02 obesity
138	117-26-3042	black	12/1/1964	female	02 chest pain
138	127-91-4819	white	10/23/1964	male	02 short of breath
139	270-89-1234	white	3/15/1965	female	02 hypertension
139	021-45-7854	white	8/13/1964	male	02 obesity
139	021-08-2839	white	5/5/1964	male	02 fever
138	117-61-0504	white	2/13/1967	male	02 vomiting
138	021-668-9440	white	3/21/1967	male	02 back pain

***k*-anonymity**



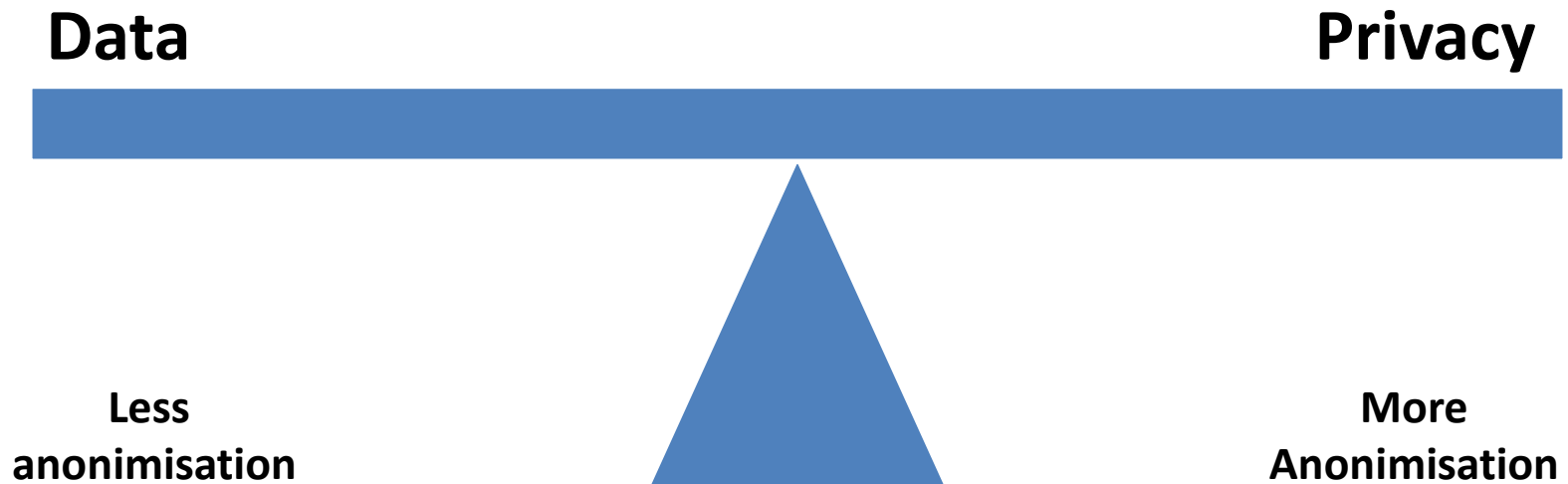
Race	BirthDate	Gender	ZIP	Problem
black	1965	male	02141	short of breath
black	1965	male	02141	chest pain
person	1965	female	0213*	painful eye
person	1965	female	0213*	wheezing
black	1964	female	02138	obesity
black	1964	female	02138	chest pain
white	1964	male	0213*	short of breath
person	1965	female	0213*	hypertension
white	1964	male	0213*	obesity
white	1964	male	0213*	fever
white	1967	male	02138	vomiting
white	1967	male	02138	back pain

Hogere *k*-waarde =>  
meer privacy en  
meer verlies van data

### Geavanceerdere aanpakken

- *l*-diversity
- *t*-closeness

# Anonimisatietechnieken



De werkelijkheid is iets complexer...

# Gefaalde anonimisatie



- 2006: publicatie 20 miljoen geanonimiseerde zoekopdrachten (650.000 users, 3 maand).
- The New York Times achterhaalde identiteit van meerdere gebruikers
- Ontslag CTO, #57 in CNNs *“101 Dumbest Moments in Business”*

The image shows the Netflix logo, which consists of the word 'NETFLIX' in a bold, white, sans-serif font. The letters are slightly shadowed to give a 3D effect. The logo is set against a red background.

- 2007: Publicatie geanonimiseerde records met filmratings 500.000 gebruikers
- Identificatie gebruikers in combinatie met publieke IMDB data
- Voor de rechter door gebruikers

# Paul Ohm

## *Broken Promises of Privacy*



*“De robuuste anonimisatieveronderstelling is niet fundamenteel foutief, maar wel diep gebrekkig.”*

*“We kunnen niet voorspellen tot welke en tot hoeveel externe informatie de aanvaller toegang heeft.”*

*“Persoonsgegevens zijn een steeds groeiende categorie. Tien jaar terug beschouwde bijna niemand filmbeoordelingen als persoonsgegevens.”*

*“Het aanwasprobleem: Eénmaal een aanvaller twee ‘geanonimiseerde’ databases aan elkaar gelinkt heeft, kan er makkelijker andere informatie aan gelinkt worden, wat kan helpen bij deanonimisatie.”*

# Whitehouse.org



“Anonimisatie van een data record lijkt misschien makkelijk te implementeren. Helaas is het steeds makkelijker om anonimisatie teniet te doen met behulp van de technieken die ontwikkeld worden voor legitieme toepassingen van big data.”

*“Sommige oudere technologieën, zoals anonimisatietechnieken, hebben maar een beperkt toekomstig potentieel, hoewel ze in het verleden wel waardevol waren.”*



*Ik ben bezorgd om security en privacy. En dat is waarom ik mezelf niet langer moet wijsmaken dat ik tot Generation Y behoor.*

— Kristof Van der Stadt

## Anonimiteit bij big data-analyse een illusie

02/02/15 om 11:19 - Bijgewerkt om 12:46

Bron: Datanews

**Big data waar persoonsgegevens uit verwijderd zijn, lijken de privacy te waarborgen. Maar er is maar weinig nodig om persoonlijke gegevens weer aan de persoon te koppelen, blijkt uit onderzoek.**

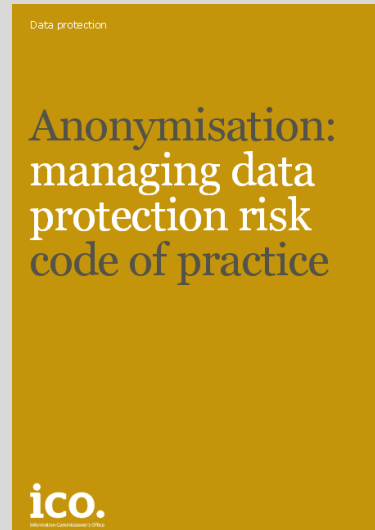
53

KEER GEDEELD





## Uitgebreide documentatie



## UK Anonymisation Network (UKAN)

*“Het is opgezet als een middel om best practices in anonimisatie toe te passen. Het biedt praktisch advies en informatie.”*



<http://ukanon.net/>

# Verenigd Koninkrijk



*“Big data is geen spel dat met andere regels gespeeld wordt.”*

*“Anonimisatie stelt organisaties in staat om zekerheid te geven aan de personen over wie het data verzamelde; [zekerheid] dat voor big data analyse geen data gebruikt wordt die hen kan identificeren.”*

*“De reeks beschikbare data sets en de kracht van big data analytics maken anonimisatie moeilijker [...] maar dit betekent niet dat anonimisatie onmogelijk is of dat het geen effectieve tool is.”*



**Ann Cavoukian**

## *Big Data and Innovation, Setting the Record Straight: De-identification Does work*

*“Er is een neiging bij een deel van de commentatoren om de bevindingen te overdrijven.”*

*“Er duikt literatuur op die de nauwkeurigheid en legitimiteit van gepubliceerde re-identificatieaanvallen in vraag stelt “*

*“Organisaties moeten een risicobeoordling doen, die rekening houdt met de huidige state of the art in de-identificatie technieken en re-identificatie aanvallen.”*

DOSSIER ARCHIEF

# Google loopt deur plat bij Witte Huis

26-03-15, 07.06u - VK

LEES LATER ★



President Obama en vicepresident Joe Biden nodigden in 2013 de toplui van Amerika's belangrijkste technologiebedrijven uit. ©GETTY

1

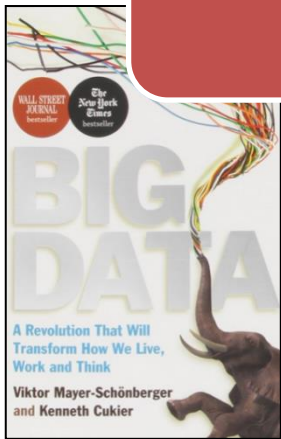
No



Yes



Wie heeft gelijk?



Mayer-Schonberger

Cukier



Populaire pers



PRINCETON UNIVERSITY



Ann Cavoukian

# Beperkingen technieken anonimisatie



# Onderbouwing

**Impliciete: close-world assumption**

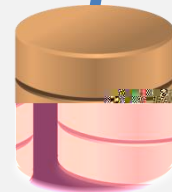
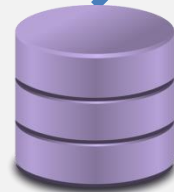
Aanvaller geen toegang tot externe data

**Wat weet aanvaller?**

- Data
- Algoritmes
- Vandaag, morgen, overmorgen, ...

**Toevertrouwen aan derden / publiceren  
“geanonimiseerde” data kan riskant zijn**

# Wat weet de aanvaller?



**Eigen  
Data/kennis**

**Gekochte  
Data**

**Gestolen  
Data**

**Publieke  
Data**

**P(Deanonimisatie) ↗**



# Aanvallen



## Large-scale

- Deanonimiseer zo veel mogelijk
- Vb. Hackers



## Targetted

- Identificeer gericht specifiek individu
- Vb. Politieke concurrent



## Opportunistisch

- Identificeer een gekende (kennis, bekende)
- Vb. Nieuwsgierige buur

# Hoogdimensionale data

## Eéndimensionale data

Uit te drukken in één getal (Dimensie)

Vb. Postcode, geboortedatum, geslacht, salaris, lengte, BMI, ...



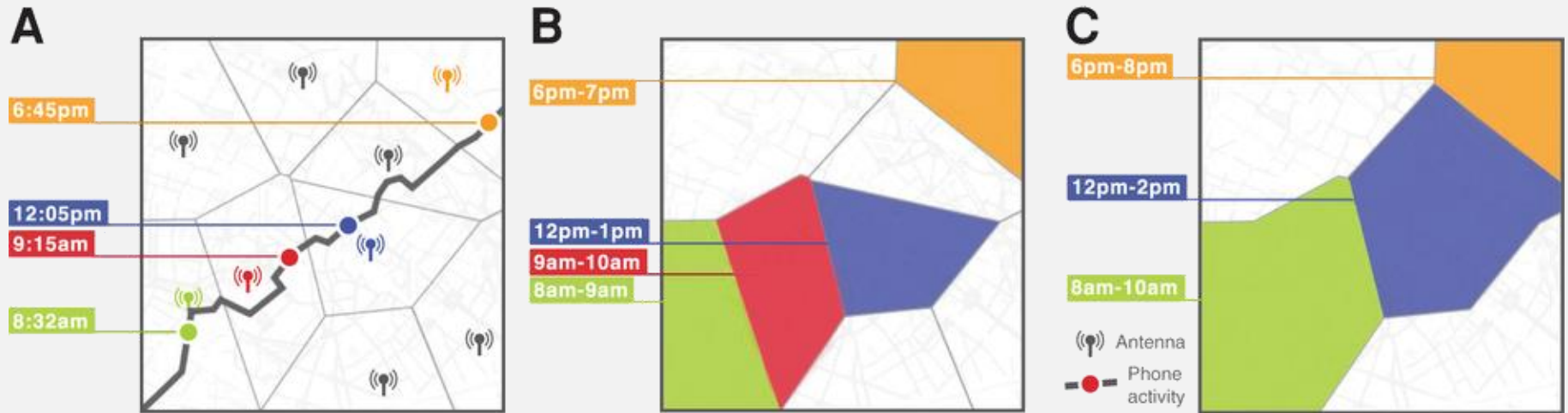
## Hoogdimensionale data

Enkel uit te drukken met een groot (vaak stijgend) aantal getallen

Vb. location tracking, friendship graph, aankoopgedrag, ...

# Hoogdimensionale data

## Locatie tracking

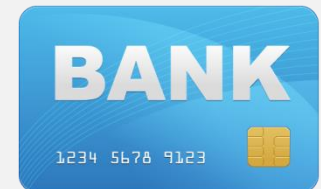


Vervagen data (tijd, locatie) heeft weinig impact op uniekheid [M13]  
Maar maakt dat wel veel minder bruikbaar

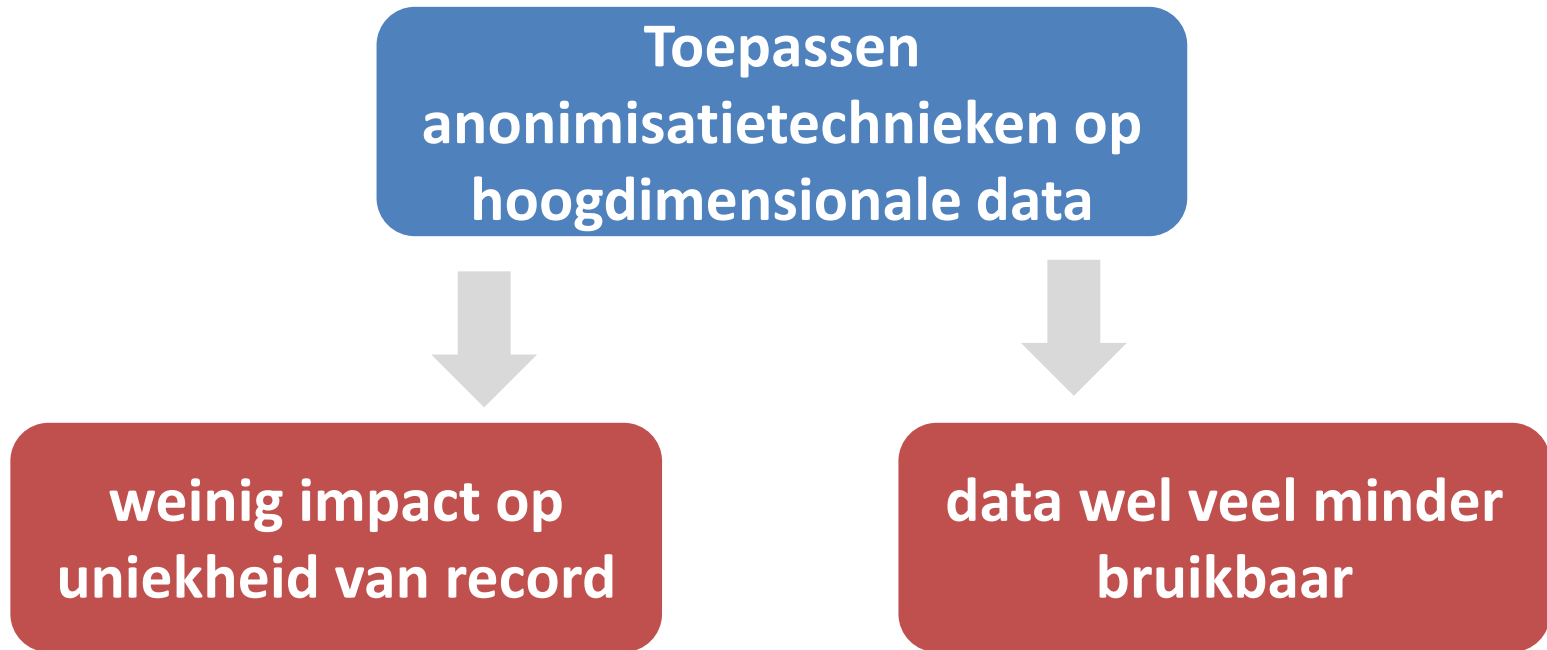
## Aankoopgedrag

4 x locatie-dag: uniek voor 90% kopers





Vrij weinig informatie al uniek



# Hoogdimensionale data



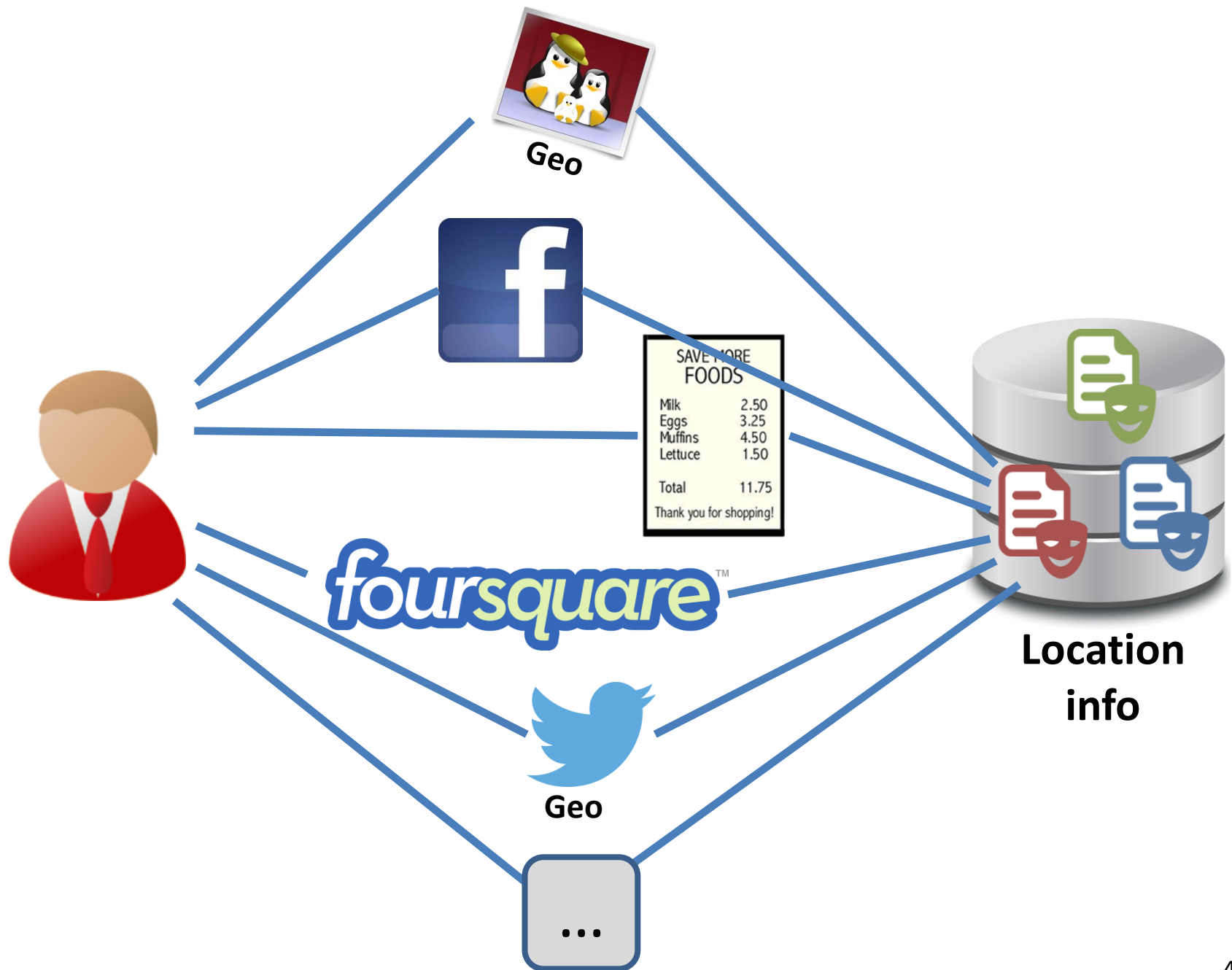
# Hoogdimensionale data

Type	Voorbeeld	Dimensies
<b>Sociale netwerken</b> 	Gemiddeld 100 vrienden op facebook	100
<b>Reviews</b> 	Gemiddelde Netflix gebruiker in dataset: 213 ratings	426 (tijd+rating)
<b>Locatie tracking</b> 	Elke 2 uur locatie + tijd	720 per maand
<b>Aankoopgedrag</b> 	4 keer per maand 20 producten aankopen	160 per maand
<b>Surfgedrag</b>	...	...
<b>Medisch dossier</b>	...	...
<b>Genetische data</b>	...	1.000.000
	...	...

Uniek

≠

Linkbaar aan individu





**Ann Cavoukian**

*Big Data and Innovation, Setting  
the Record Straight: De-  
identification Does work*

*“In het geval van hoogdimensionale data kunnen  
bijkomende regelingen nagestreefd worden, zoals  
het ter beschikking stellen van de data aan  
onderzoekers onder strikte voorwaarden.”*

**Slechts één zin in haar tekst...**

**(waarin ze toegeeft dat deanonimisatietechnieken niet steeds werken)**

# Hoogdimensionale data

De norm in big data

Balans privacy en nut  
uit evenwicht in anonimisatie-  
paradigma

**We kunnen data niet en juridisch geanonimiseerd maken  
en tegelijkertijd waardevol houden voor analytics doeleinden**

(Wel nog potentieel in meer klassieke context)

# Conclusie

Geen wetenschappelijke onderbouwing in open world

Meer data beschikbaar / Grotere datasets

Hoogdimensionale data niet te anonimiseren

Verschillende types aanvallers

Anonimisatietechnieken op zich onvoldoende  
voor analytics

Privacywet blijft van toepassing

# Kruisen van Persoonsgegevens

## Een fictief voorbeeld

*Een onderzoeksteam wil medische, financiële en demografische persoonsgegevens analyseren van alle burgers die na 1990 geboren zijn, die zelfstandige in bijberoep zijn met een loon van minstens € 50 000.*

*Deze gegevens worden echter beheerd door verschillende overheidsbedrijven en moeten dus gekruist worden.*

Kan technologie steeds persoonsgegevens converteren naar anonieme gegevens die vervolgens geanalyseerd kunnen worden?



Wat is er technisch mogelijk binnen het wettelijke kader?

# AGGENDA

Introductie



Technieken Anonimisatie

Wat

Beperkingen

Data Archipel

Concept



Proof of concept

Een beetje crypto

Deanonimisatie

Sleutelbeheer

Transparantie

Small cells



Conclusies

# Data Archipel

© Smals Research

# Valida

**THE WALL STREET JOURNAL.**  
**In Belgium, an Encryption Powerhouse Rises**  
University of Leuven has become a battleground in the fight between privacy and surveillance



## COSIC, KU Leuven

- Technische validatie
- *“Zit conceptueel mooi in elkaar”*



## Dept. Computer Science, KU Leuven

- Technische validatie
- *“Zeer interessant”, “Zeer actueel probleem”*
- Samenwerking wetenschappelijk artikel



## Prof. dr. Patrick Van Eecke (UA, DLA Piper)

- Juridische validatie
- *“Geen showstoppers”, “een stap vooruit”, “hoogstwaarschijnlijk interesse bij privé”*

# Wetenschappelijke Publicatie

Wetenschappelijk artikel  
*The Data Archipelago -  
Reconciling privacy and  
analytics on multi-source PII*  
i.s.m. dept. computerwet. KU Leuven

?

Conference  
*16th Privacy Enhancing  
Technologies Symposium  
(PETS 2016)*

Journal  
*Proceedings on Privacy  
Enhancing Technologies*

## Top crypto and security conferences

Conference	CIF	AR	PR	CR
1. IEEE S&P	3.83	12.2%	9.2%	4.7%
2. Usenix Sec	3.40	16.1%	8.1%	5.2%
3. Eurocrypt	2.92	20.3%	9.9%	4.1%
4. Crypto	2.63	21.1%	12%	4.9%
5. NDSS	2.40	17.1%	20.8%	3.7%
6. ACM CCS	2.36	18.4%	15.6%	8.3%
7. CHES	2.35	25.6%	9.2%	7.7%
8. Asiacrypt	2.25	16.9%	19.1%	8.5%
<b>9. PETS</b>	<b>2.16</b>	<b>22.9%</b>	<b>14.1%</b>	<b>9.4%</b>
10. ACSAC	1.94	21.1%	18%	12.4%

CIF =  $1 / (AR+PR+CR)$ , where

AR = No. accepted papers / No. of submissions

PR = No. accepted papers / No. of registered participants

CR = No. accepted papers / No. of citations

# Kruisen van Persoonsgegevens

## Een fictief voorbeeld

*Een **onderzoeksteam** wil medische, financiële en demografische persoonsgegevens analyseren van alle **burgers** die na 1990 geboren zijn, die zelfstandige in bijberoep zijn met een loon van minstens € 50 000.*

*Deze gegevens worden echter beheerd door verschillende **overheidsbedrijven** en moeten dus gekruist worden.*

**Wetenschapper**  
Vlot kruisen data

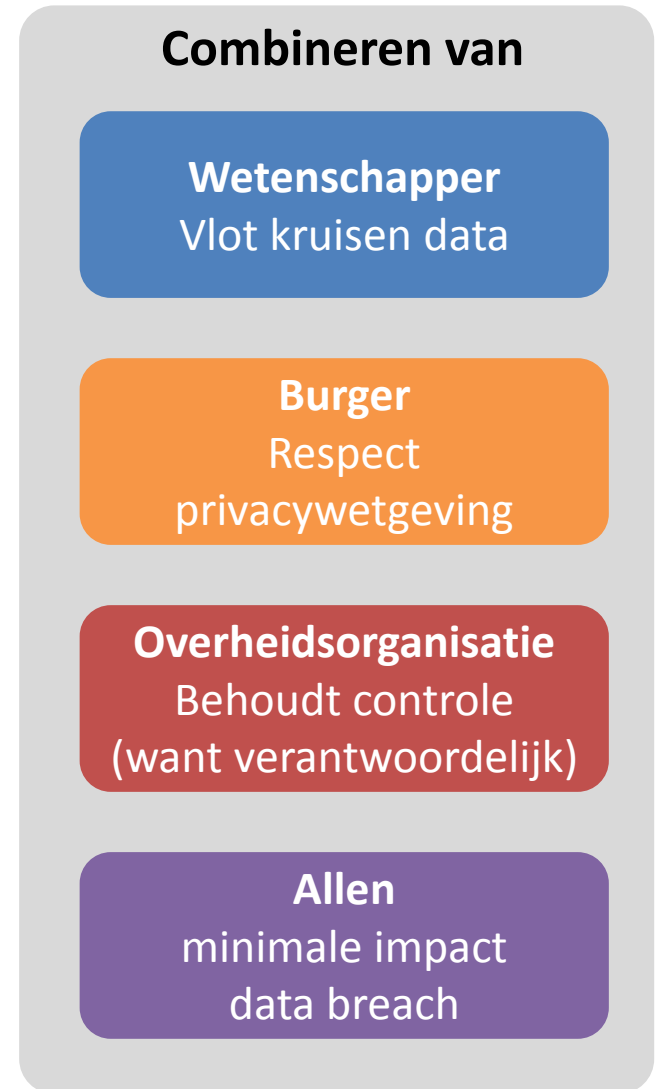
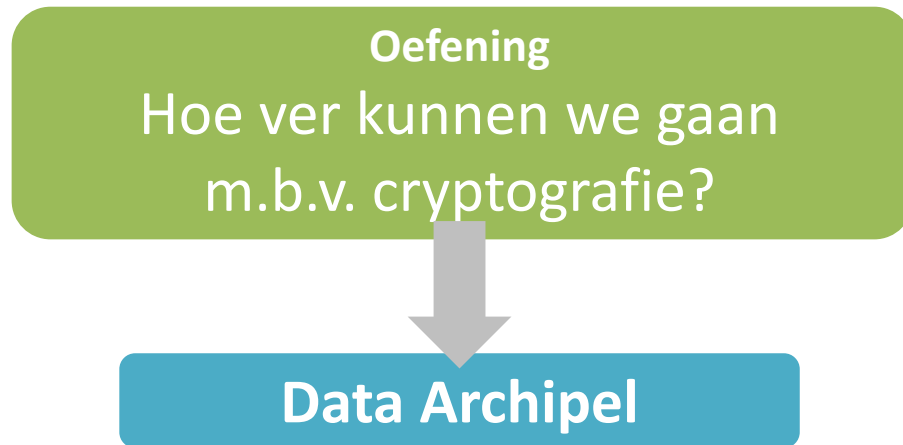
**Burger**  
Respect  
privacywetgeving

**Overheidsorganisatie**  
Behoudt controle  
(want verantwoordelijk)

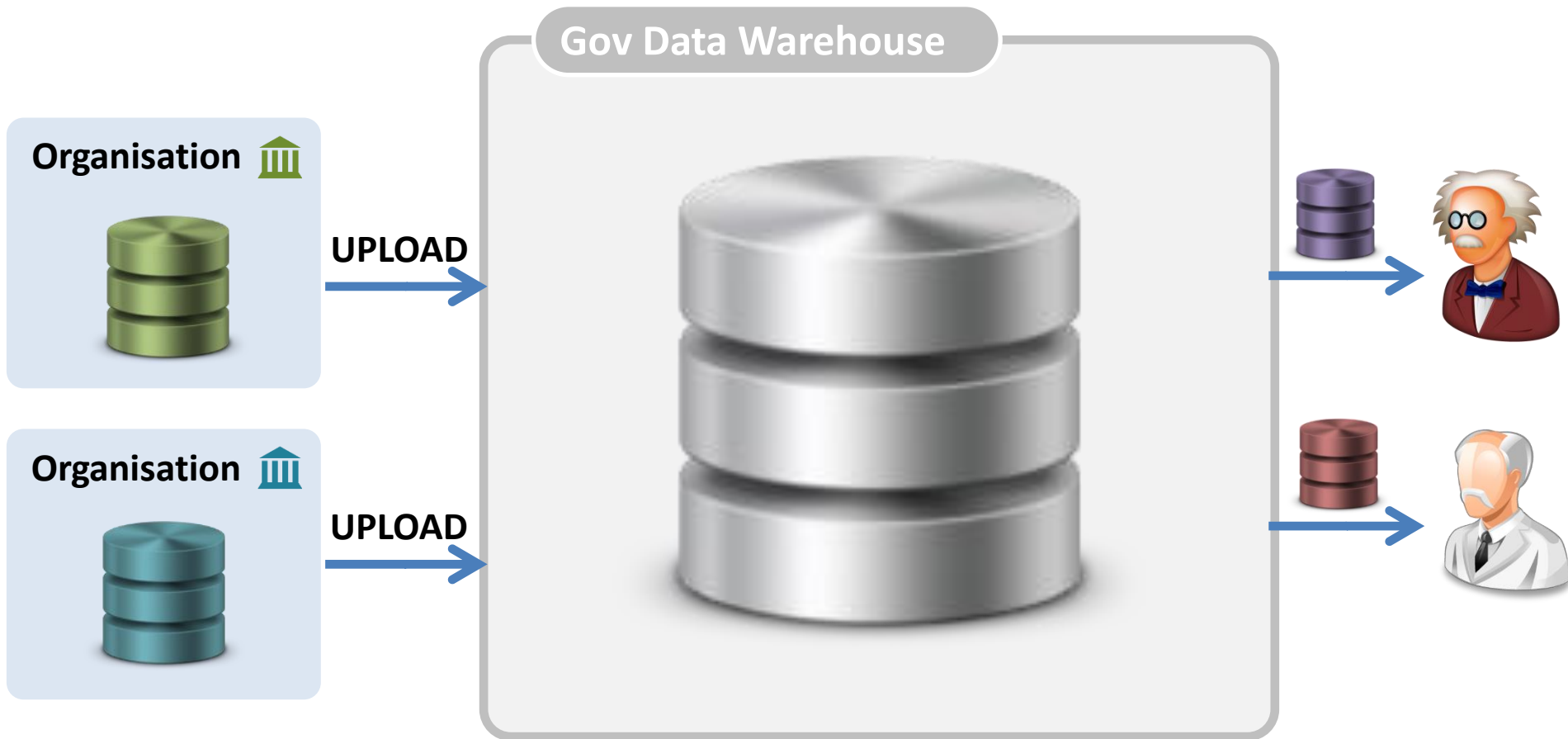
**Allen**  
minimale impact  
data breach

# Kruisen van Persoonsgegevens

Een fictief voorbeeld



# Digital Data Dystopia



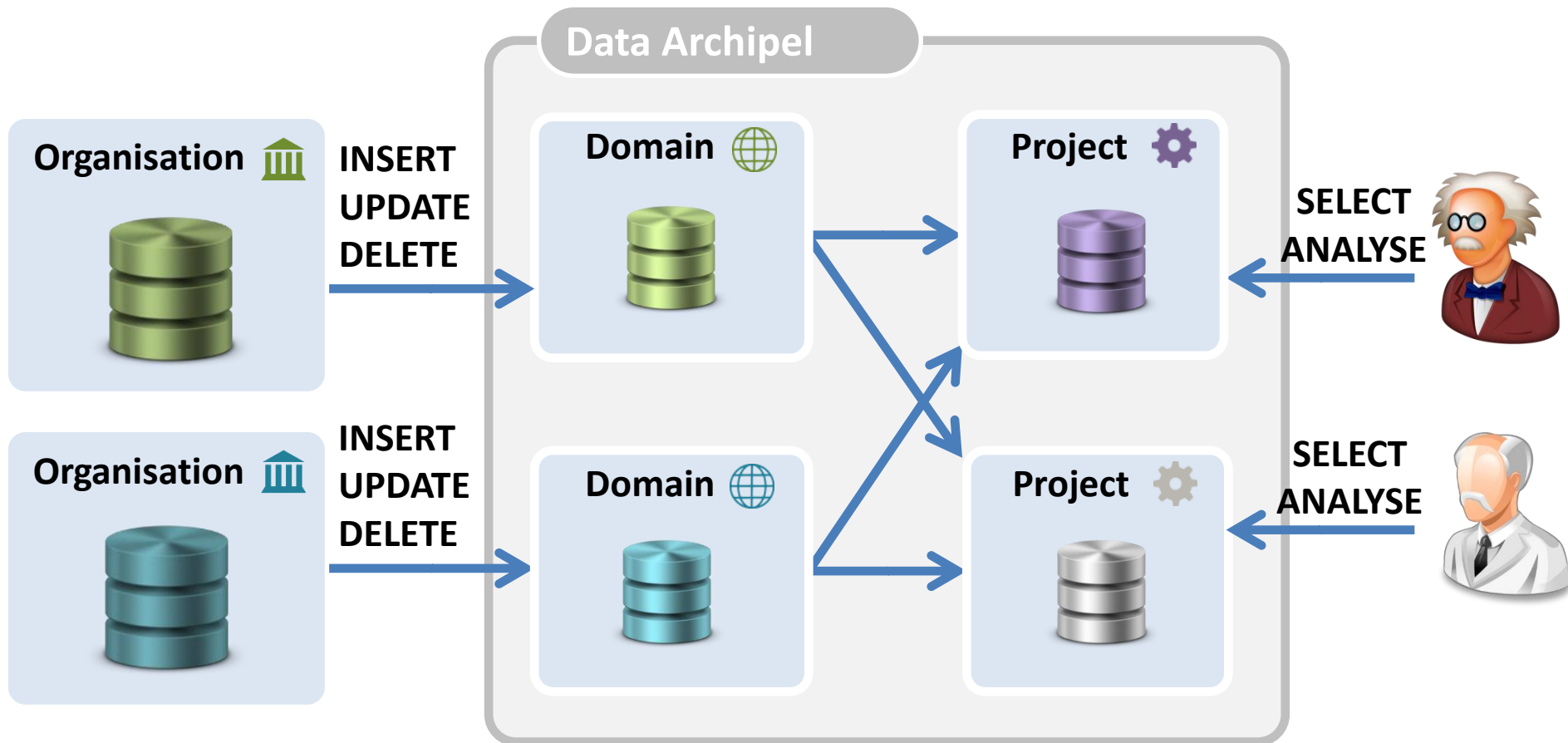
Overheidsorganisatie  
geen enkele controle

Data breach  
dramatisch

Privacy  
risico's

Vlot Kruisen  
data

# Concept



## Domain

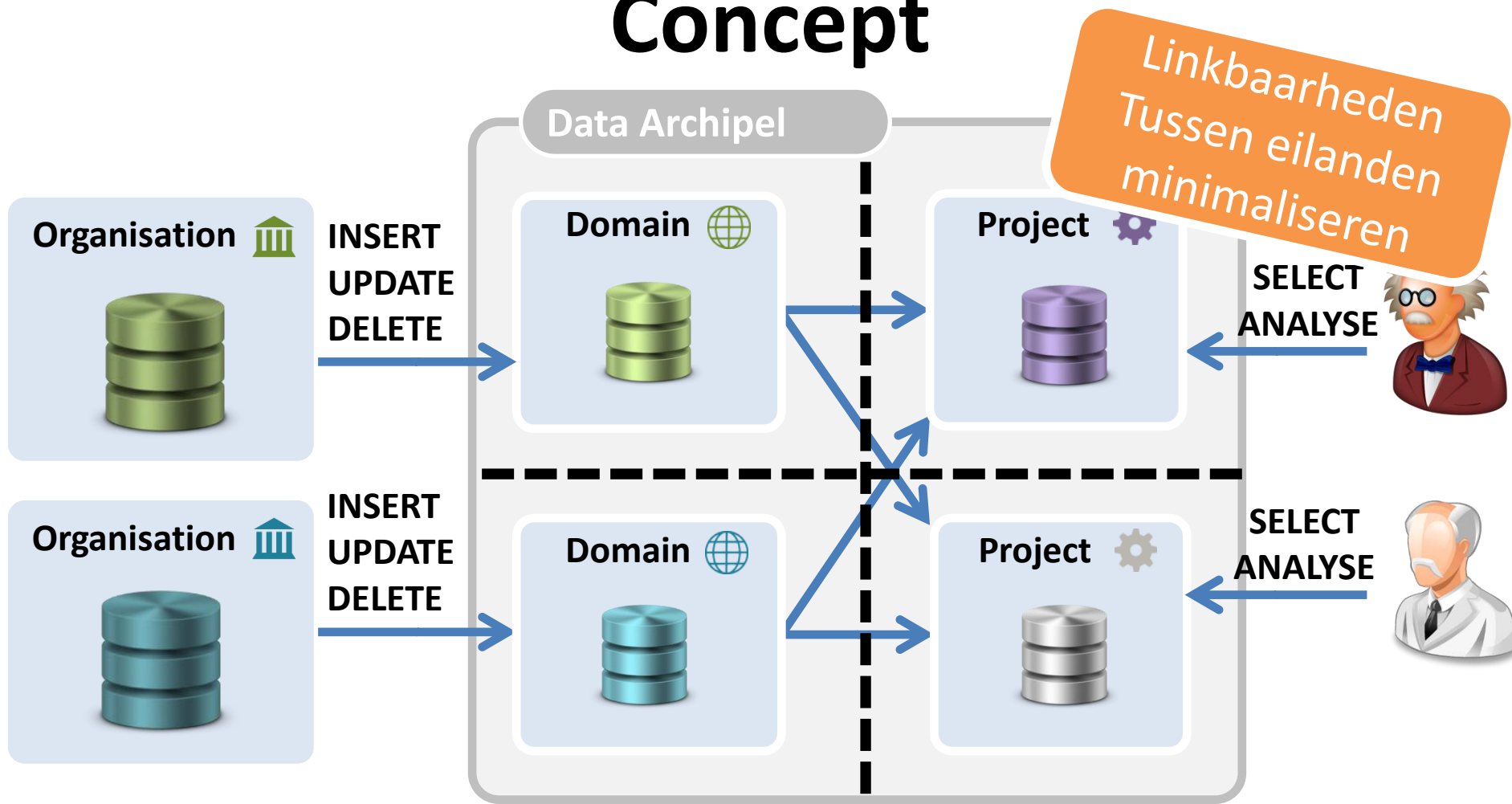
- Beheerd & gecontroleerd door één organisatie
- Permanent
- Lage performantievereisten



## Project

- Ontvang minimaal vereiste data
- Toegangscontrole + monitoring
- Tijdelijk
- Hoog performant

# Concept



Maximale controle  
overheidsorganisatie

Kleinere impact  
bij data breach




Mechanismes  
Privacy




Vlot kruisen  
data




# Linkbaarheid

**Voorkom**

Data Model


Domain   
 € 

Domain   
 € 


Project   
 € 

**Elimineer**

Data Model



Domain   
 € 



Project   
 € 



Project   
 € 

**Minimaliseer**

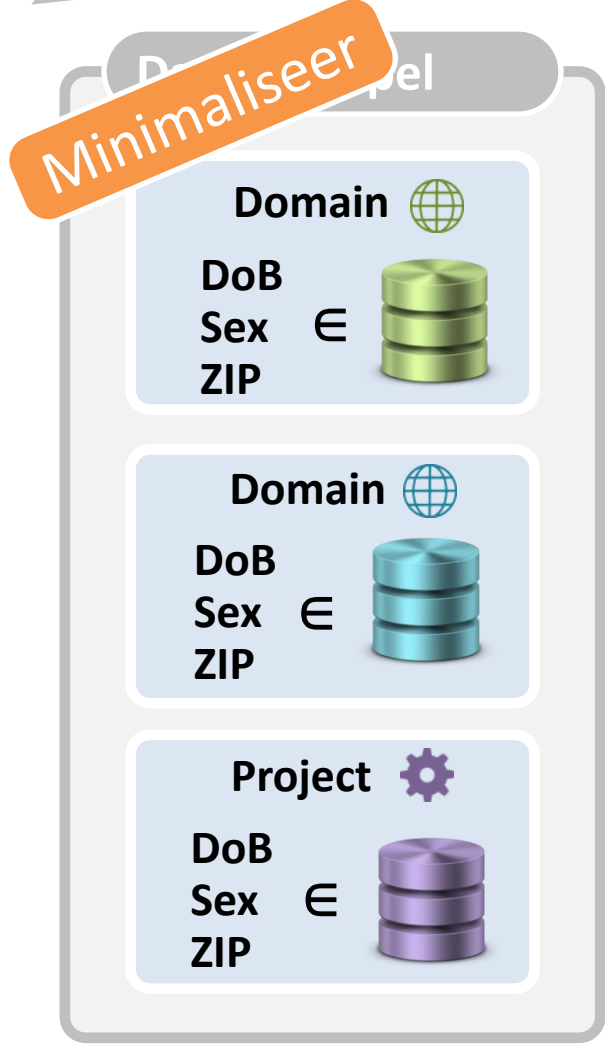
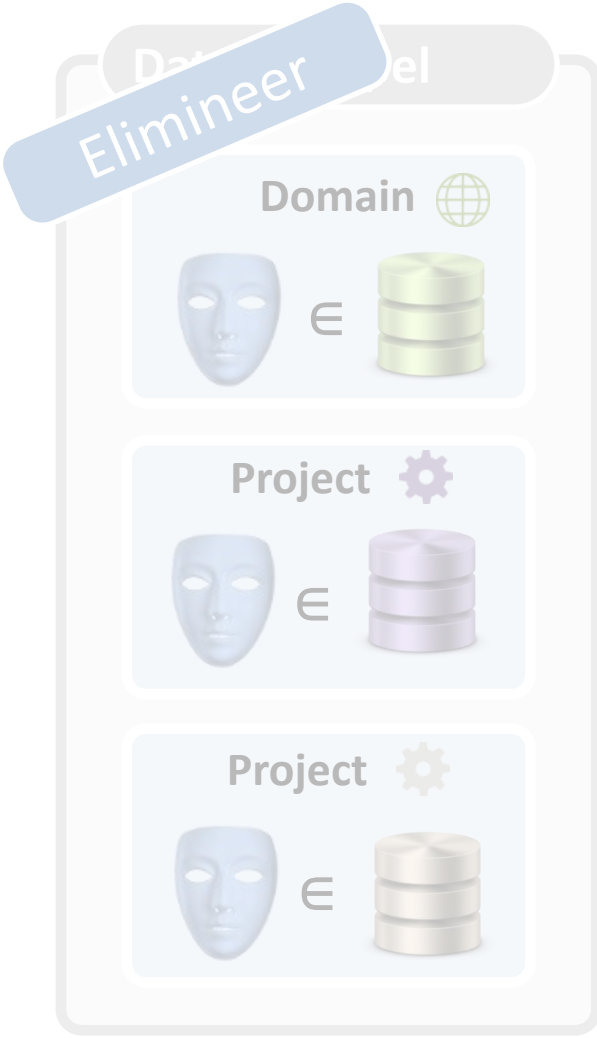
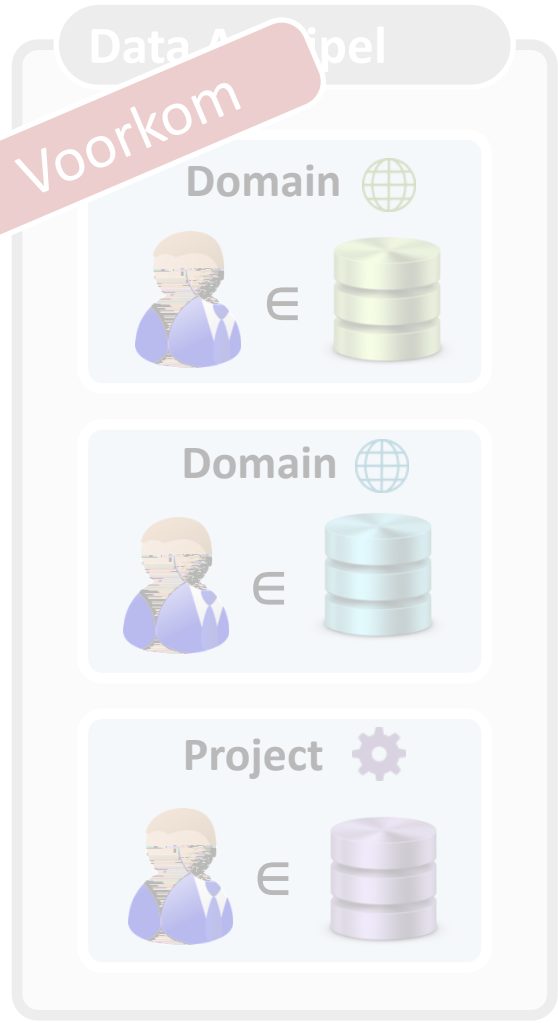
Data Model

Domain   
DoB  
Sex €   
ZIP

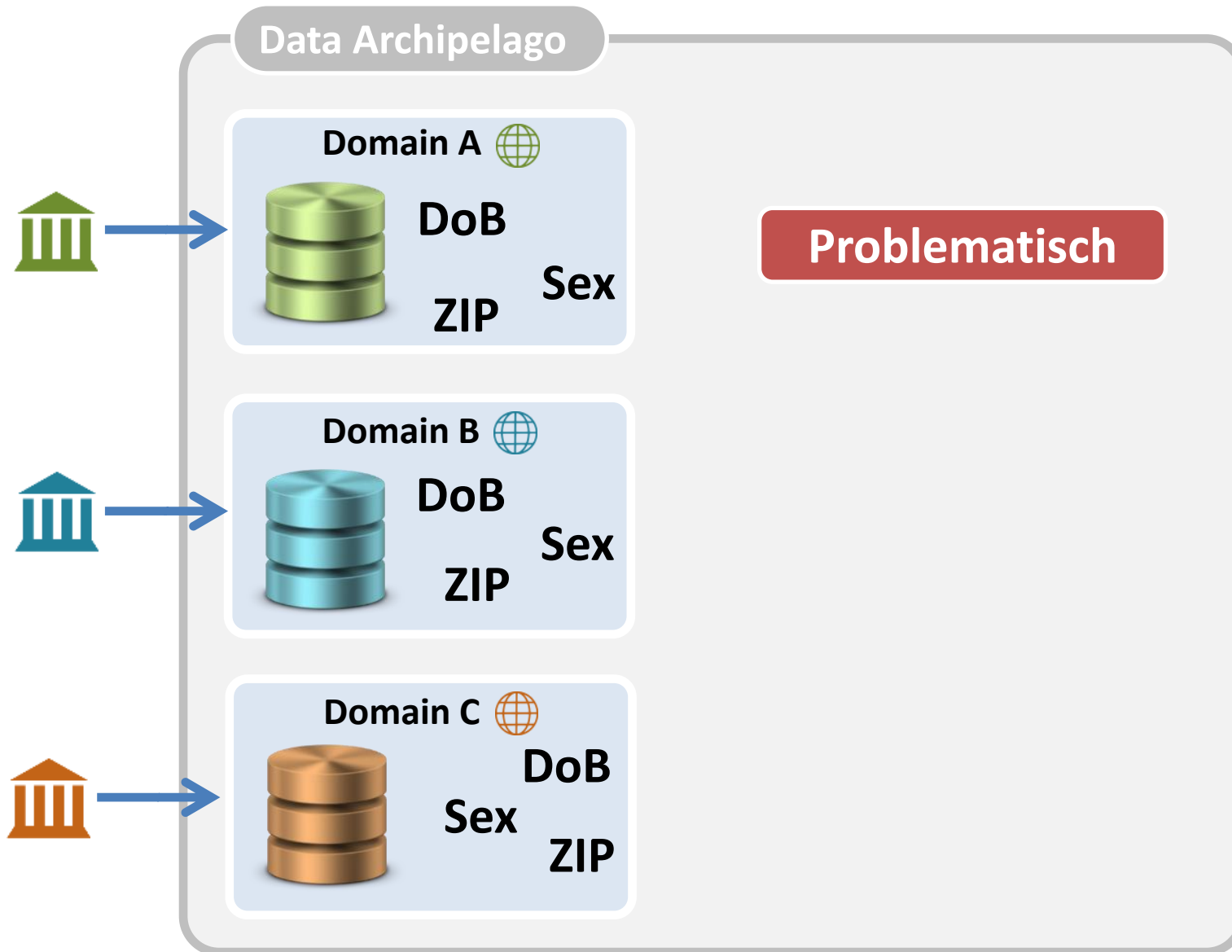
Domain   
DoB  
Sex €   
ZIP

Project   
DoB  
Sex €   
ZIP

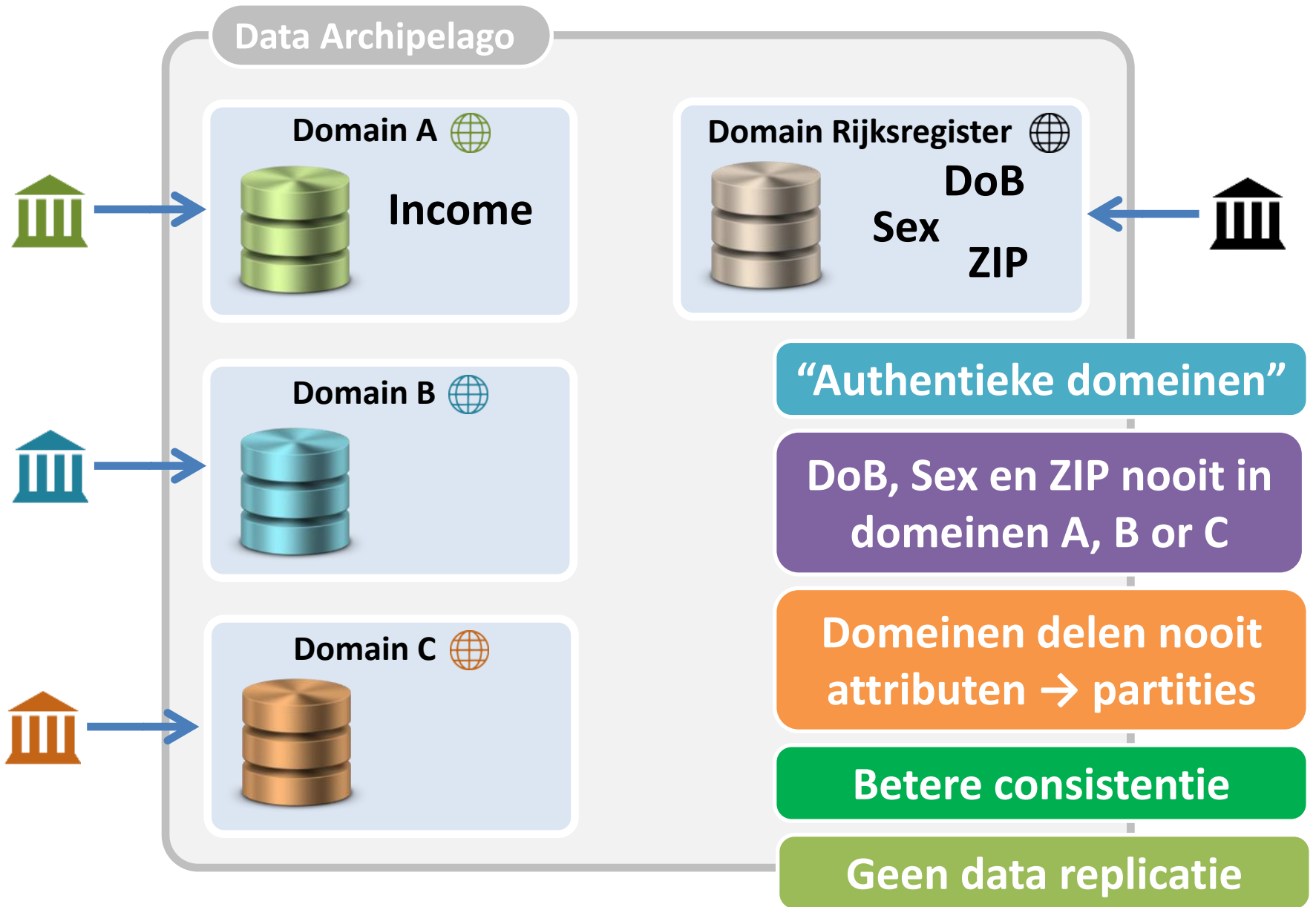
# Linkbaarheid



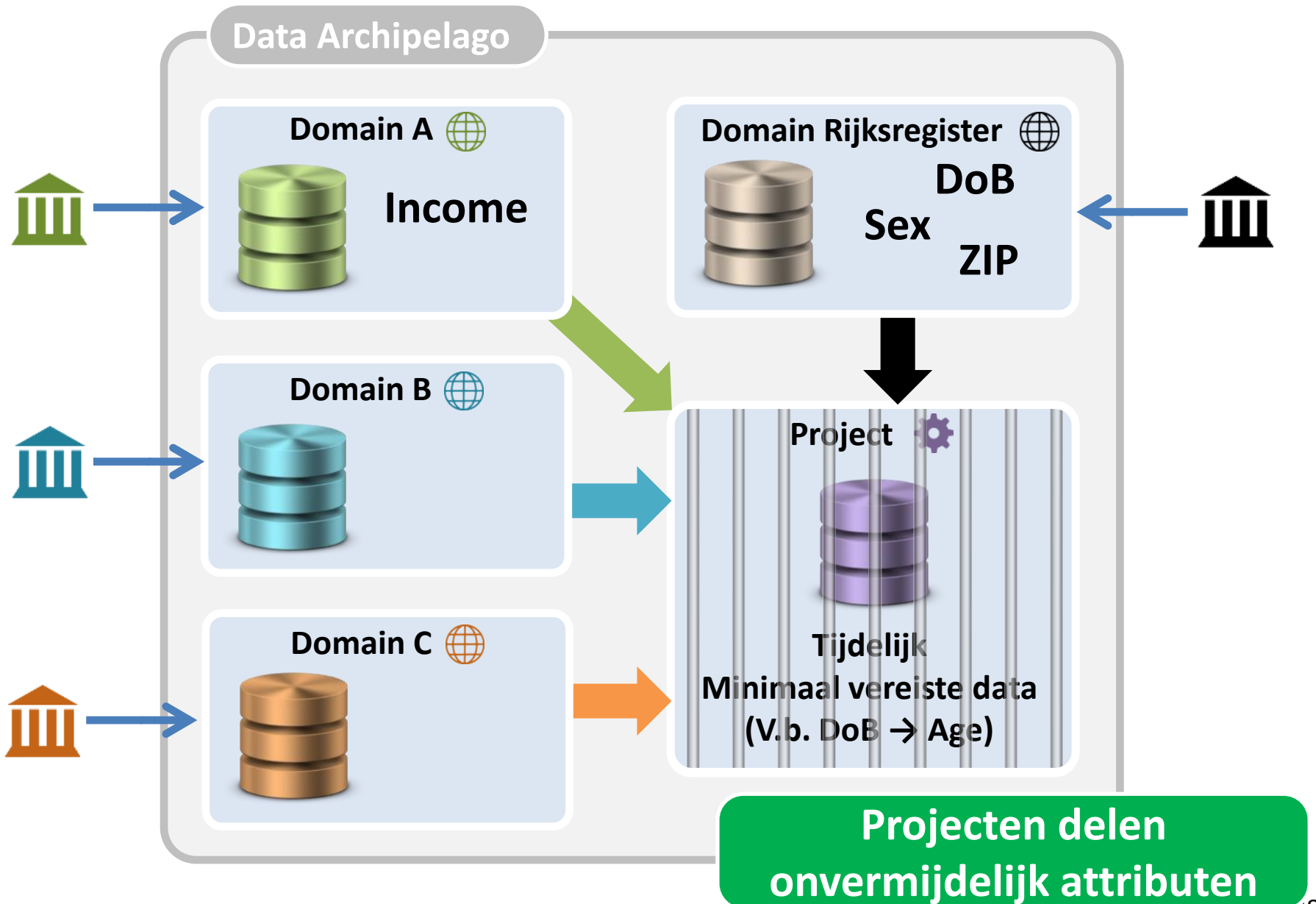
# Attribuut-linkbaarheid



# Attribuut-linkbaarheid






# Attribuut-linkbaarheid









# Linkbaarheid

**Voorkom**

Data Model




Domain   
 € 




Domain   
 € 




Project   
 € 

**Elimineer**

Data Model



Domain   
 € 



Project   
 € 



Project   
 € 

**Minimaliseer**

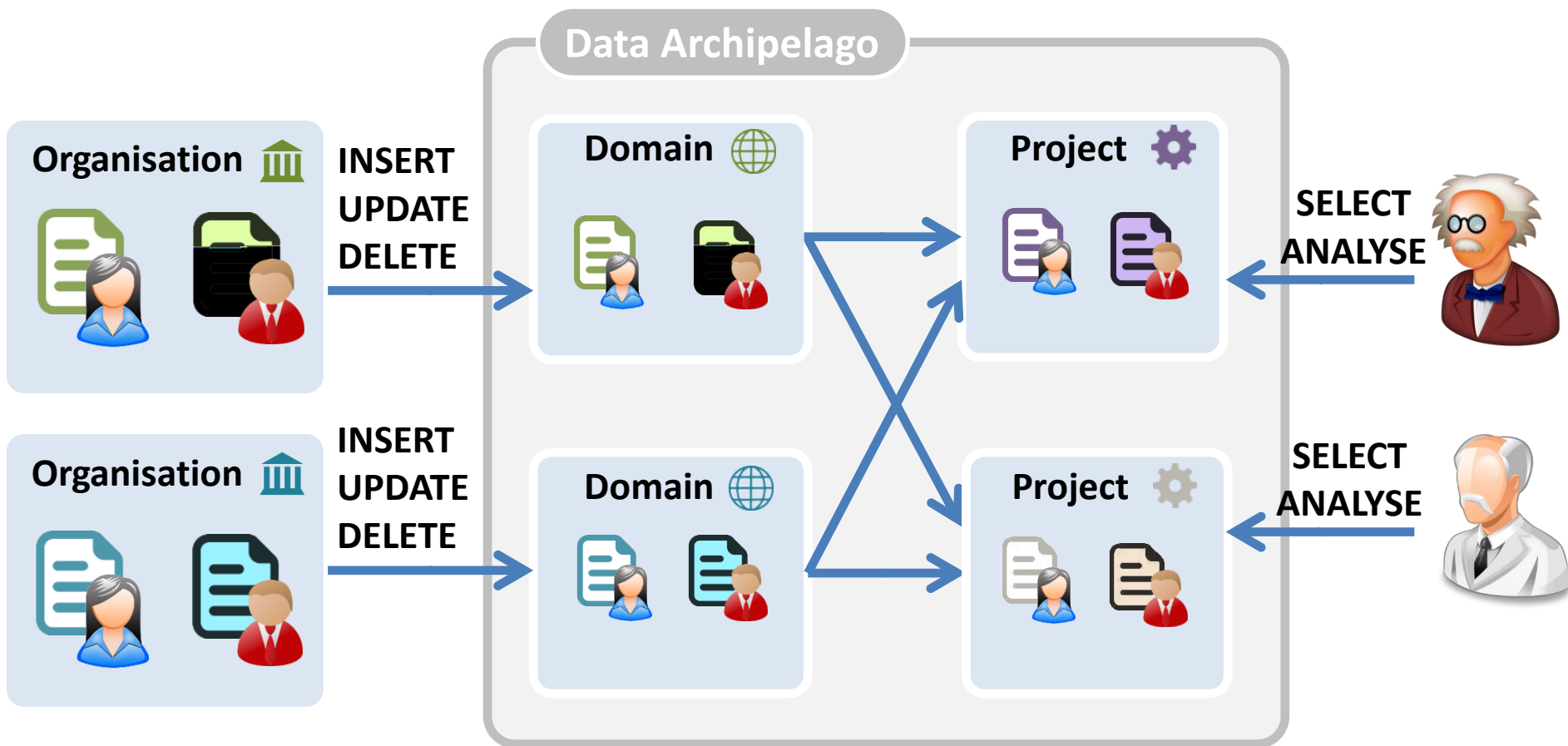
Data Model

Domain   
DoB  
Sex €   
ZIP

Domain   
DoB  
Sex €   
ZIP

Project   
DoB  
Sex €   
ZIP

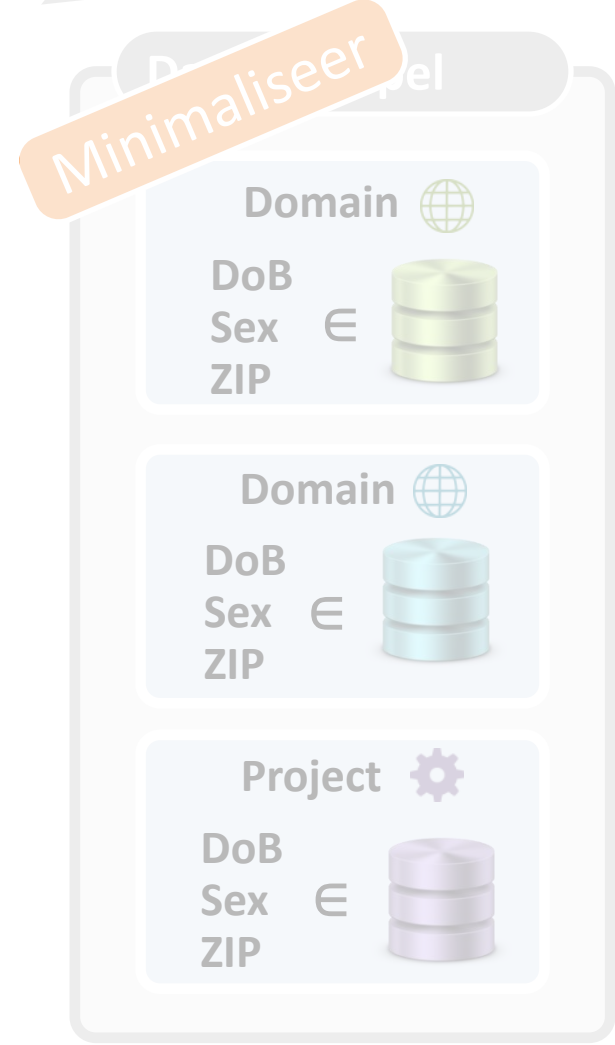
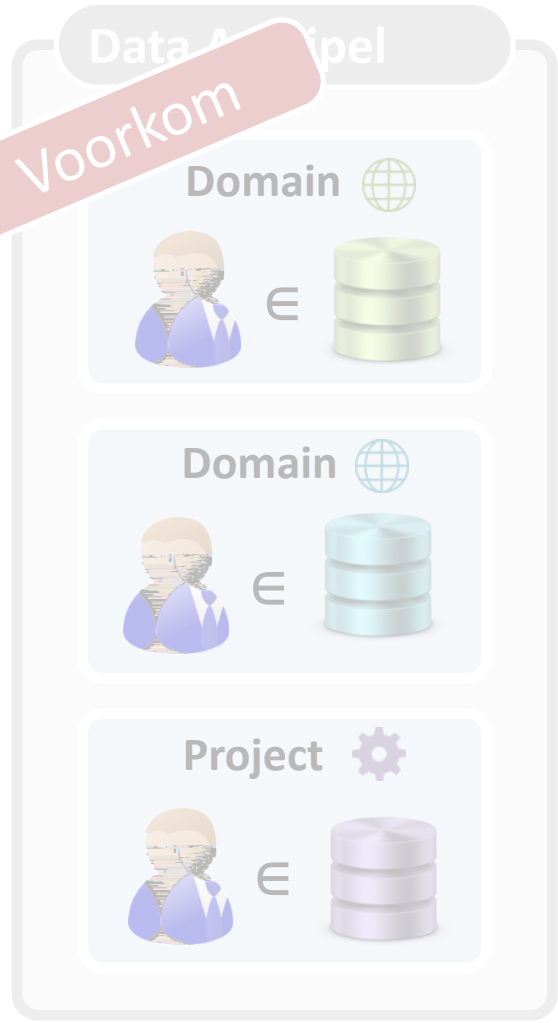
# Linkbaarheden met identifiers



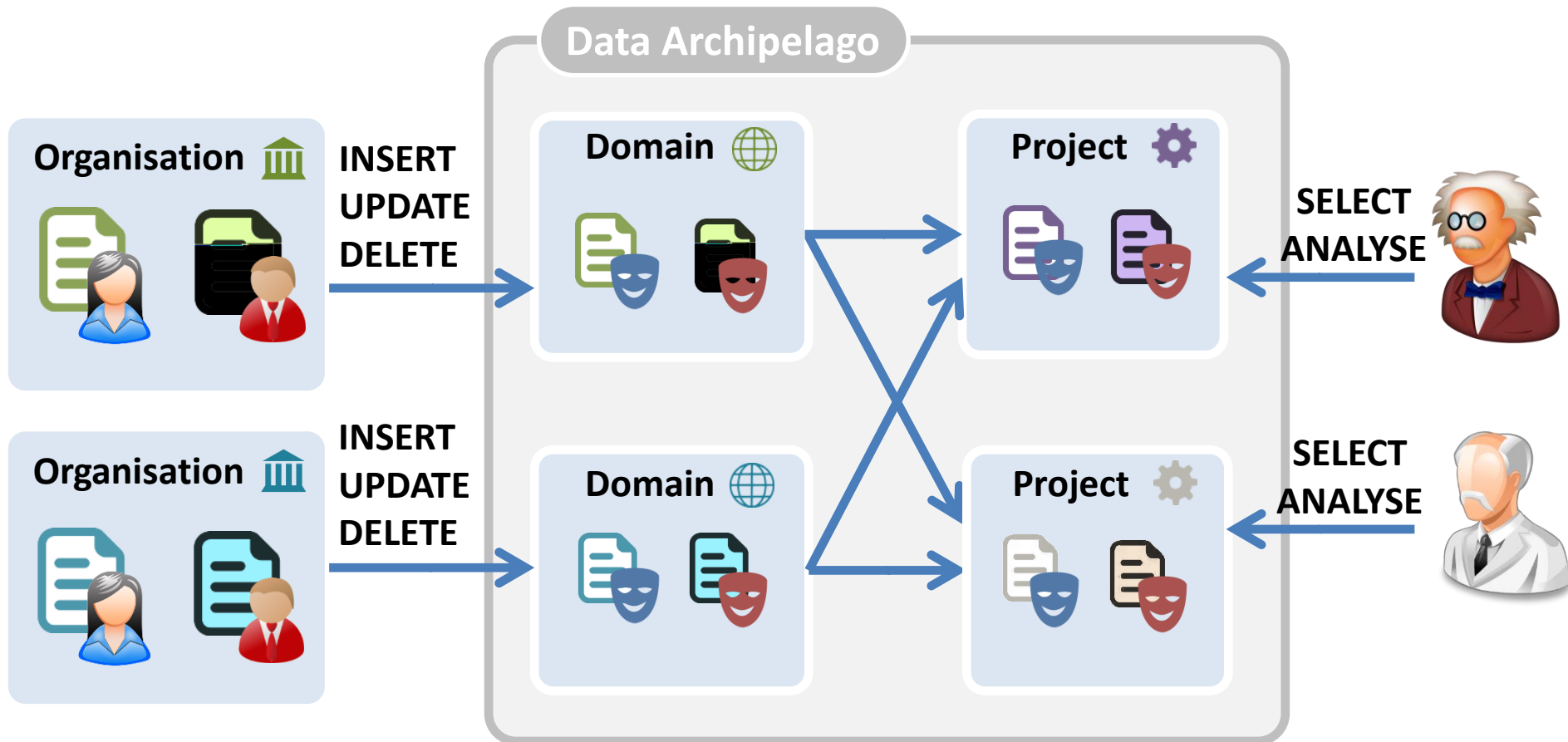
Elk eiland kent de identifier (INSZ-nummer) van burger

**We laten niet toe dat identifiers in de data archipel komen**

# Linkbaarheid

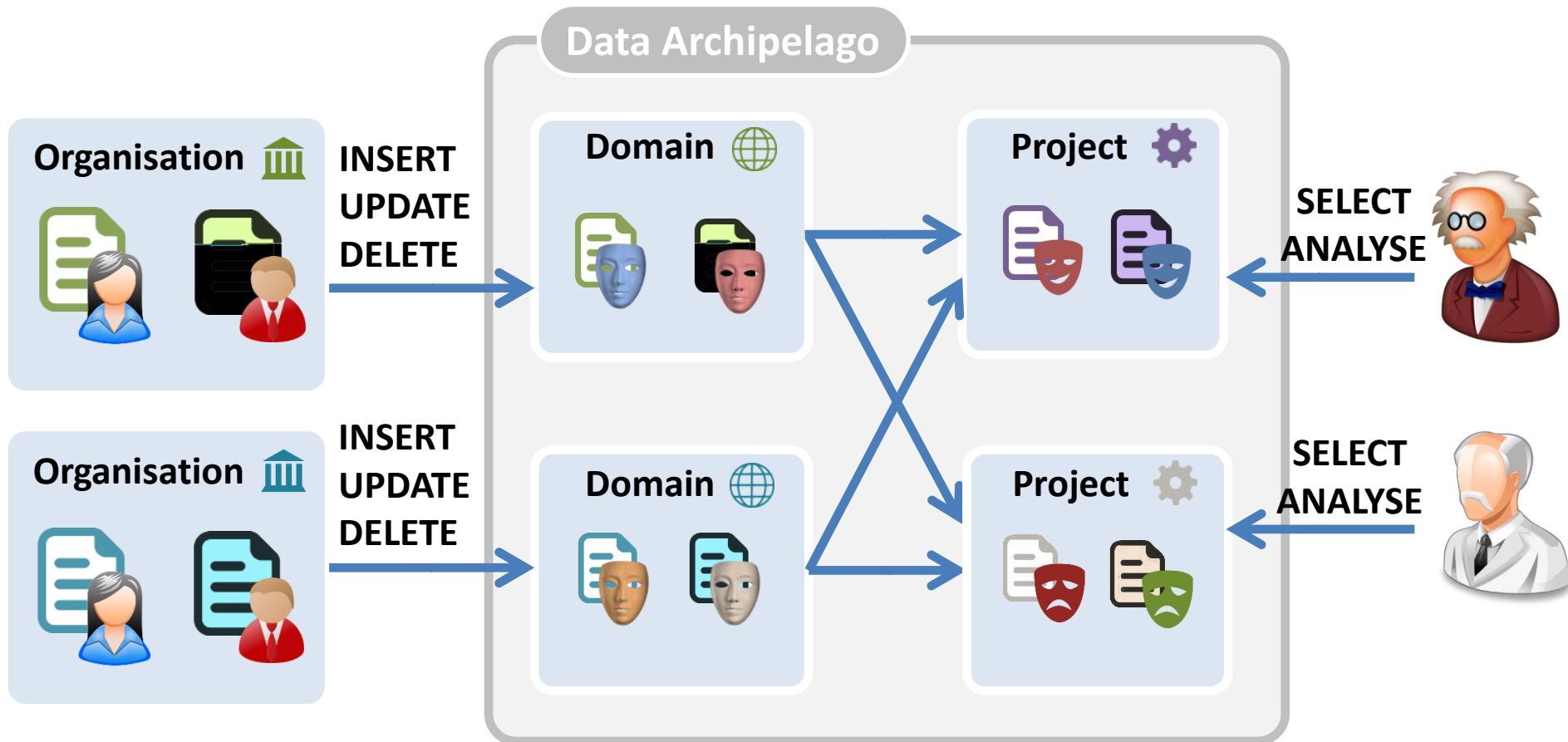


# Linkbaarheden met pseudoniemen



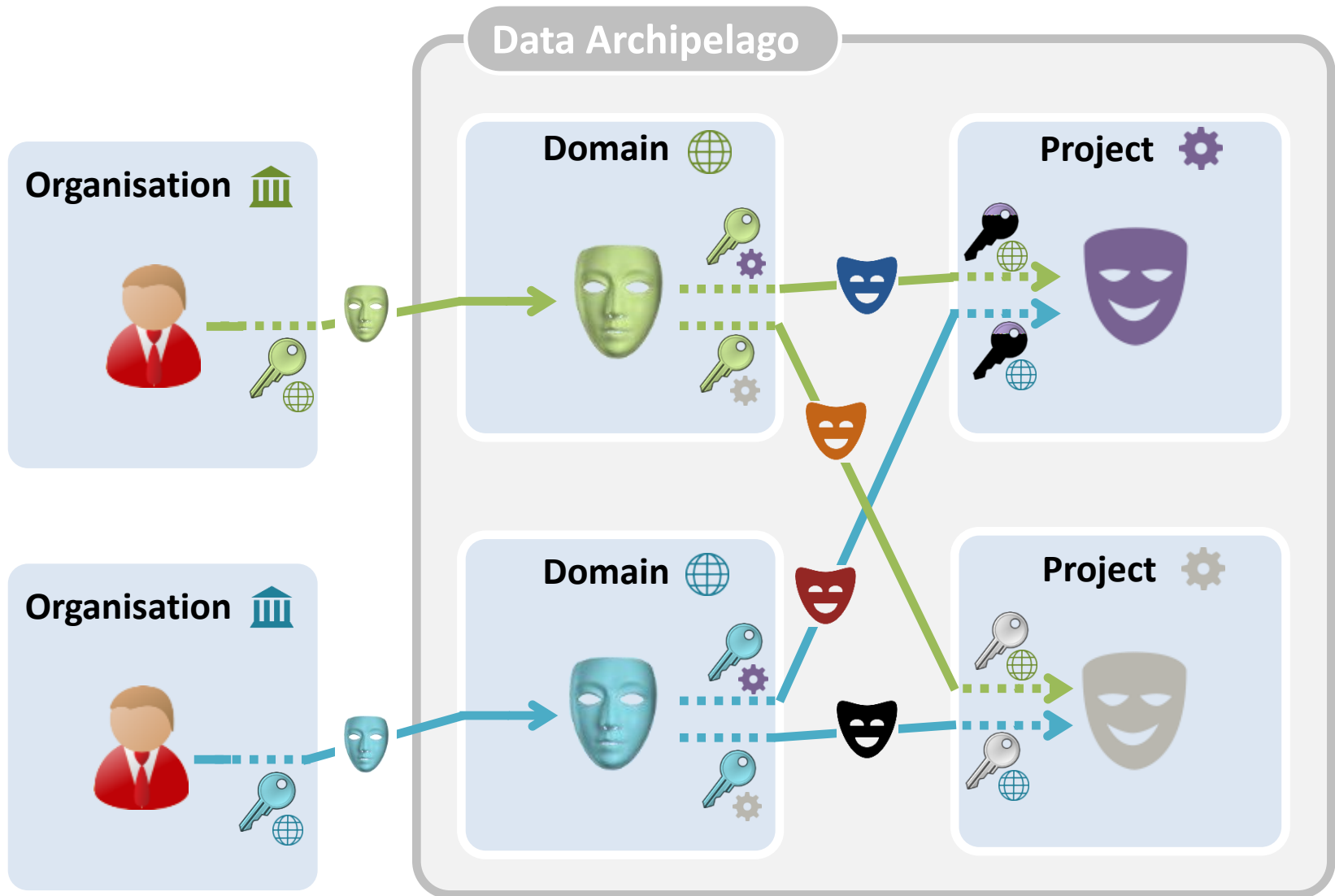
Burger elk eiland gekend onder ander pseudonym  
Pseudonymen onlinkbaar

# Ideale situatie



Elke burger heeft apart pseudonym per eiland  
Pseudoniemen onlinkbaar aan elkaar en aan identifier

# Identifiers & Pseudonymen




Nieuwe link vereist minstens twee partijen => isolatie

# Een project opstarten

Project heeft data nodig over

**Domein Rijksregister**   
Geboortejaar  $\geq$  1990

**Domein A**   
Zelfstandige  
in bijberoep

**Domein B**   
Loon  $>$   
50.000€/jaar

## Stappen


1. Goedkeuring project (machtigingsaanvraag)

2. Het vereiste sleutels worden gegenereerd

3. De vereiste data wordt door project verzameld

# Identifiers & Pseudoniemen

## Data Archipelago

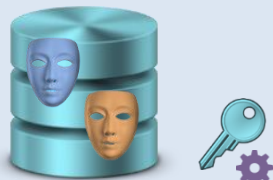
Domain RR 



Domain A 



Domain B 



Project 



# Identifiers & Pseudoniemen

## Data Archipelago

Domain RR 



“YoB  $\geq$  1990”



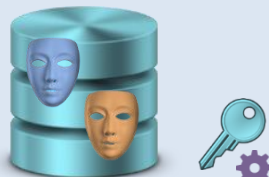
Domain A 



“zelfstandige  
in bijberoep”



Domain B 



“Loon > 50 000€”



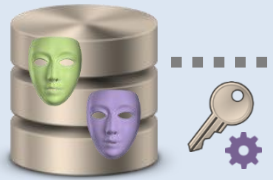
Project 



# Identifiers & Pseudoniemen


## Data Archipelago

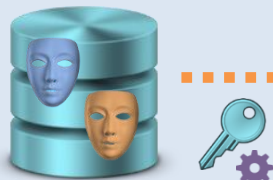
Domain RR 



Domain A 



Domain B 



Project 



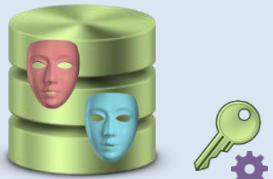
# Identifiers & Pseudoniemen

## Data Archipelago

Domain RR 



Domain A 



Domain B 



Project 



$\cap$



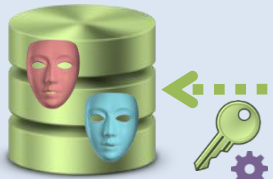
# Identifiers & Pseudoniemen

## Data Archipelago

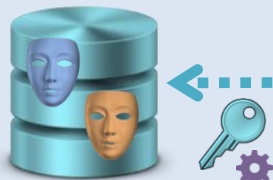
Domain RR 



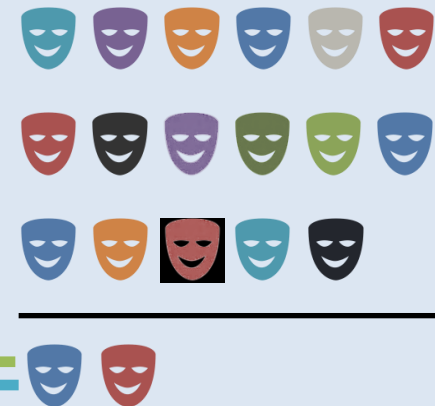
Domain A 



Domain B 



Project 



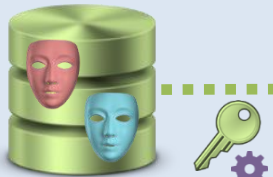
# Identifiers & Pseudoniemen

## Data Archipelago

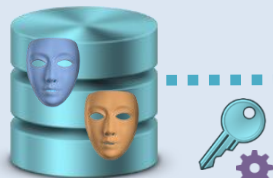
Domain RR 



Domain A 



Domain B 




Project 

Maximale controle door domeinen (organizaties)

Project ontvangt enkel minimaal vereiste data



# Misbruik

Niemand kan te weten komen dat ik data vraag over meer pseudoniemen 

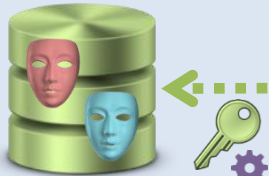
## Data Archipelago

### Domain RR

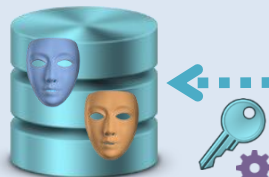


Ik leer over  en  :  
YoB  $\geq$  1990 en wage  $>$   
50.000/year 

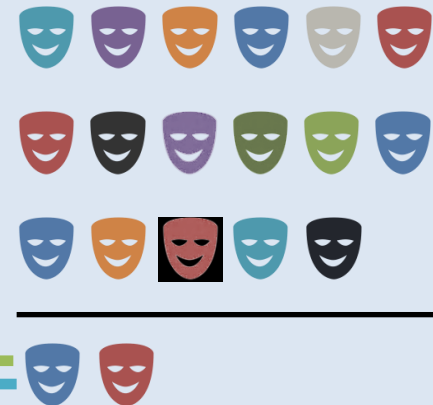
### Domain A



### Domain B

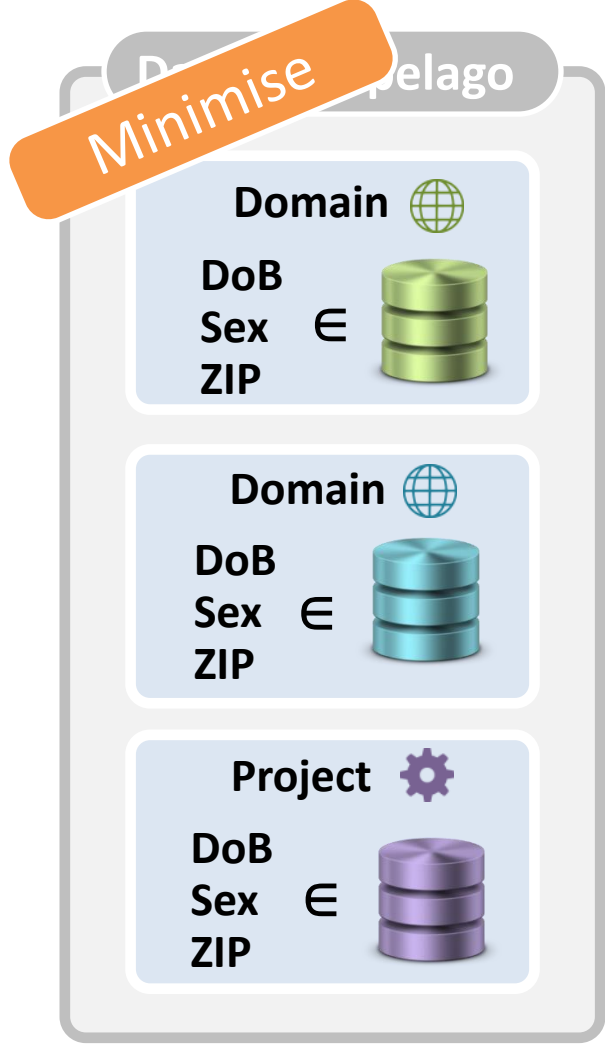


## Project

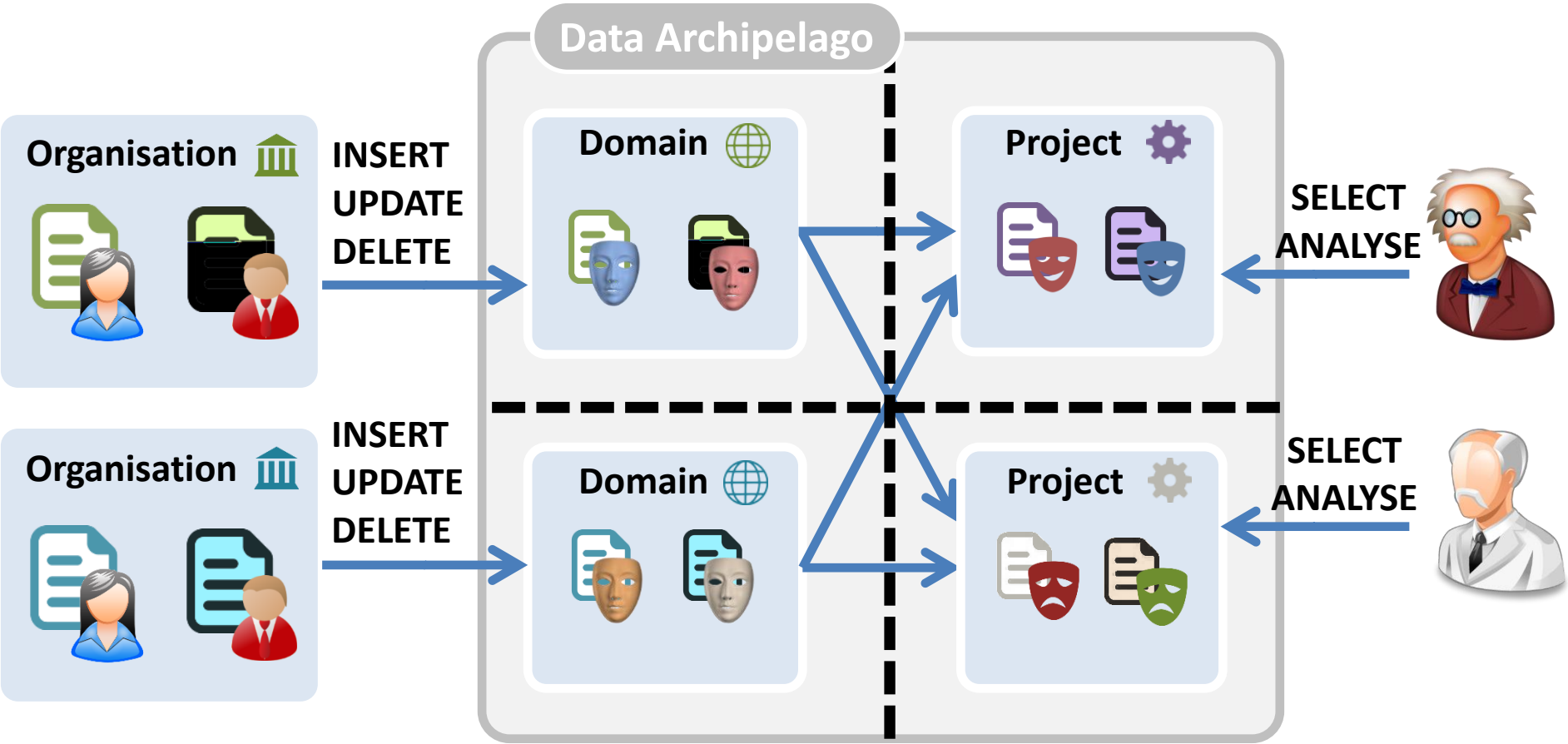


Zie wetenschappelijk artikel voor details

# Linkbaarheid



# Maximale isolatie



# Samengevat

## Organisationeel

### Organisatie behoudt controle over data

- Beslist welke data naar domain
- Medewerking vereist bij kruisen gegevens in project

### Vlot kruisen data

Medewerking betrokken organisaties vereist

## Privacy & Data Security

### Attribuut linkbaarheid geminimaliseerd

Moeilijker voor aanvaller

### Pseudoniem linkbaarheid geëlimineerd

Tenzij crypto sleutels gekend

# Pauze



# AGGENDA

- Introductie
- Technieken Anonimisatie
  - Wat
  - Beperkingen
- Data Archipel
  - Concept
  - 
  - Proof of concept
  - Een beetje crypto
  - Deanonimisatie
  - Sleutelbeheer
  - Transparantie
  - Small cells
- Conclusies



# Privacywet

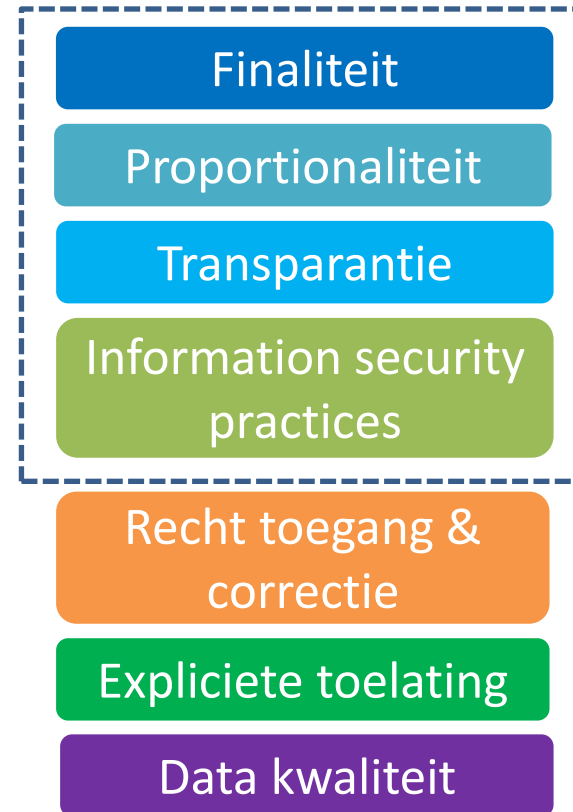
**Anonieme gegevens**

Onidentificeerbaar

**Persoonsgegevens**

Identificeerbaar/  
Geïdentificeerd

Privacywet niet  
van toepassing



**Waarschijnlijkheid op identificatie**

# Kruisen van Persoonsgegevens

## Een fictief voorbeeld

*Een onderzoeksteam wil medische, financiële en demografische persoonsgegevens analyseren van alle burgers die na 1990 geboren zijn, die zelfstandige in bijberoep zijn met een loon van minstens € 50 000.*

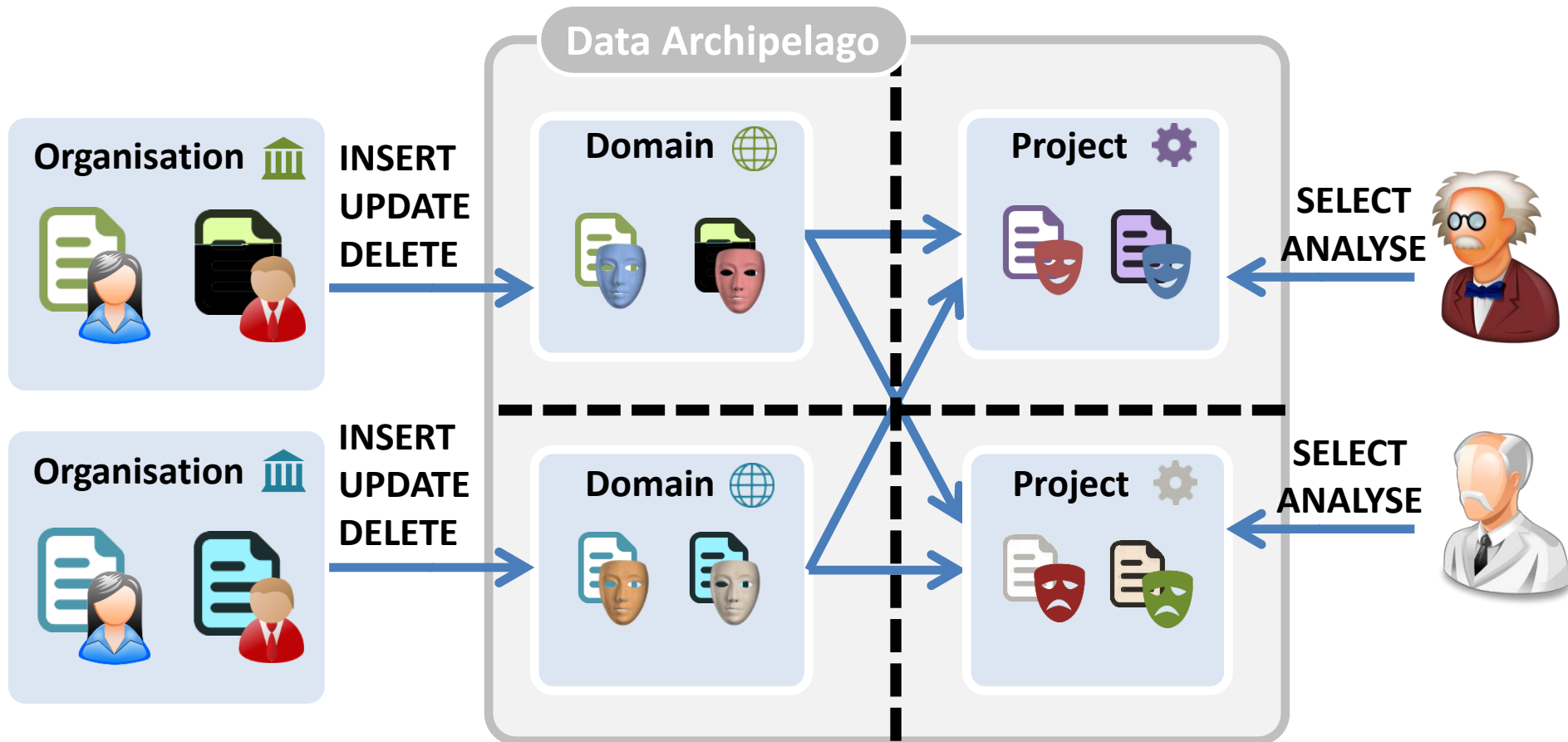
*Deze gegevens worden echter beheerd door verschillende overheidsbedrijven en moeten dus gekruist worden.*

Kan technologie steeds persoonsgegevens converteren naar anonieme gegevens die vervolgens geanalyseerd kunnen worden?

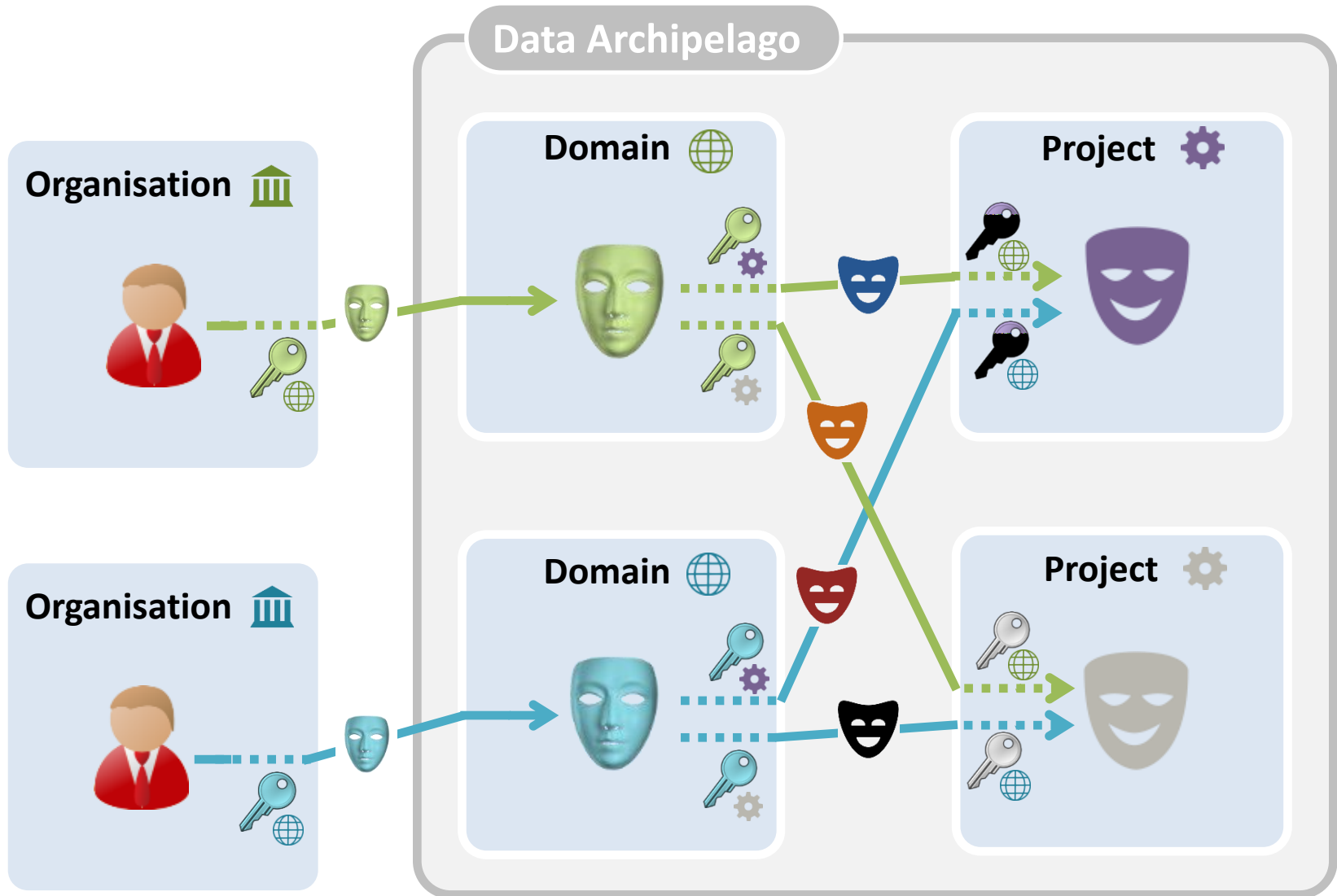


Wat is er technisch mogelijk binnen het wettelijke kader?

# Recap – Data Archipelago



# Recap – Data Archipelago



# Kruisen van Persoonsgegevens

## Een fictief voorbeeld

*Een **onderzoeksteam** wil medische, financiële en demografische persoonsgegevens analyseren van alle **burgers** die na 1990 geboren zijn, die zelfstandige in bijberoep zijn met een loon van minstens € 50 000.*

*Deze gegevens worden echter beheerd door verschillende **overheidsbedrijven** en moeten dus gekruist worden.*

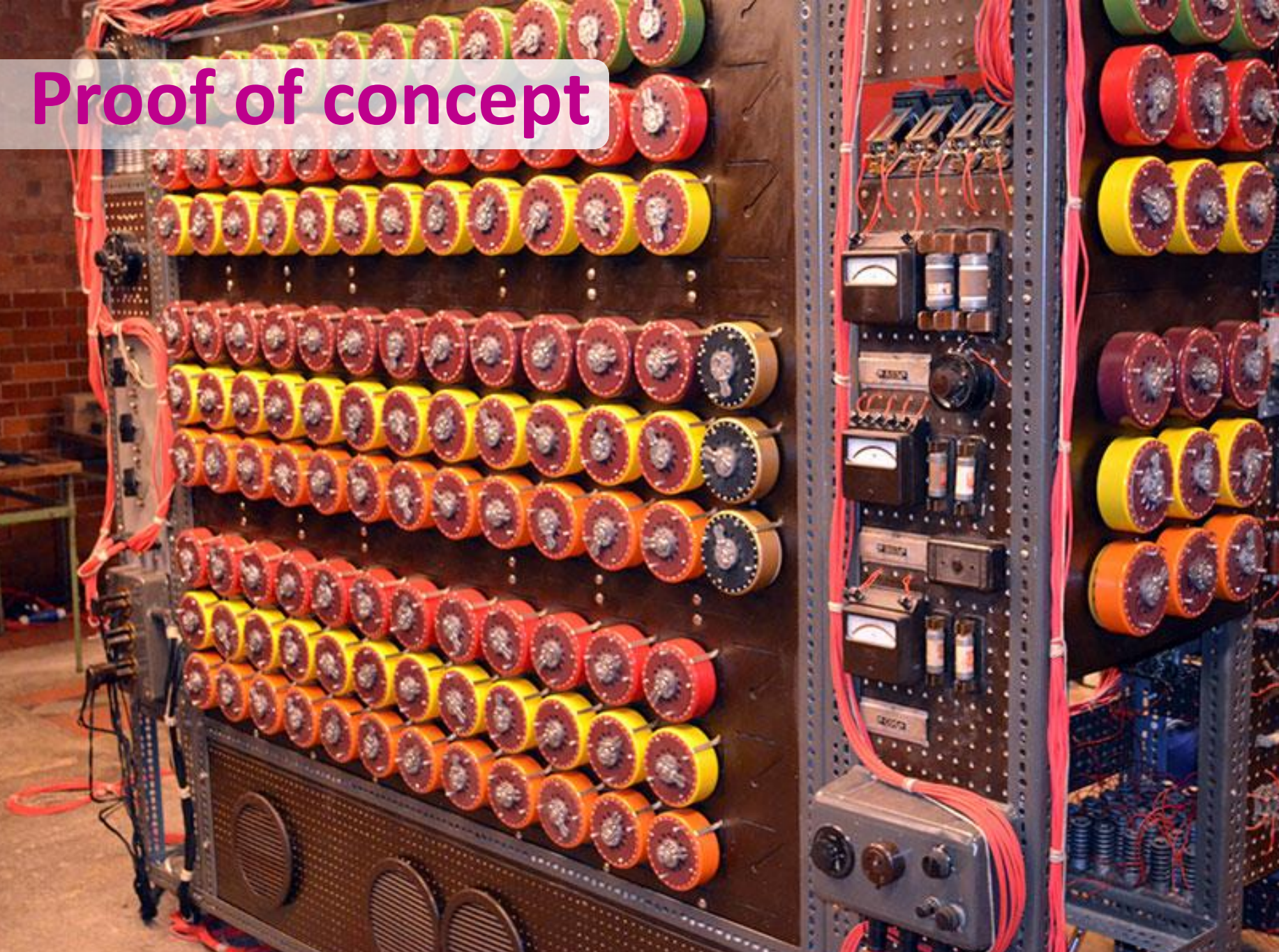
**Wetenschapper**  
Vlot kruisen data

**Burger**  
Respect  
privacywetgeving

**Overheidsorganisatie**  
Behoudt controle  
(want verantwoordelijk)

**Allen**  
minimale impact  
data breach

# Proof of concept



Execute Step Coalesce Project protocol completed: tables merged as K.U.Leuven: project Alpha's Final Data

## RijksRegister:Data

## Records

- ▶ 52041912004
- ▶ 50090272156
- ▶ 07111157873
- ▶ 00061717330
- ▶ 45051810755
- ▶ 93062824964
- ▼ \*31071211667
  - BirthDate{1931-07-12}
  - Domicile{3470;Ransberg}
  - Name{Marion Peeters}
- ▶ 92022424788
- ▶ 84122664360
- ▶ 34012980783

## RijksRegister:AnonDataDomain:...

## Records

- ▶ xebok-sydaz-kumyh-vunor-reful-guhyk-tihad-gak...
- ▶ xebom-sebop-tuneg-cemiv-mykiz-rovyg-vohod-...
- ▶ xebof-zydok-syfis-hidez-medam-mepez-pymef-b...
- ▼ \*xebom-febip-pizyl-kazyn-sidof-zefed-fygot-mod...
  - BirthDate{1931-07-12}
  - Domicile{3470;Ransberg}
- ▶ xebid-fovac-cefys-rudut-bikok-feses-resal-fuhym-...
- ▶ xebig-dihys-budot-depen-lahyc-lepyl-nyzuc-bagi...
- ▶ xebog-lumek-bezot-fagyh-kohuf-fures-bysot-coc...
- ▶ xebop-fyvus-dabak-pydab-basyh-nyfyz-lisuc-huk...
- ▶ xebip-topip-tozam-bobyf-zamam-ryper-tuhoh-n...
- ▶ xebok-vufub-dasec-sogoh-pikuc-cezil-zacoz-pyvi...

## K.U.Leuven: project Alpha's Final...

## Records

- ▶ xebol-hanor-gusot-taryg-katak-ticai-tocem-saga...
- ▶ xebol-nozas-famyn-dameg-fozic-secyt-pygot-raf...
- ▶ xebob-hubik-nevik-masab-hulek-cugin-kezo-kor...
- ▶ xebim-lyrem-kaser-lodib-bovec-gomug-gyfyz-fys...
- ▼ \*xebib-zenaf-bypac-sazyv-lukyf-gusar-kyhyl-seva...
  - BirthDate{1931-07-12}
  - Domicile{Province{Vlaams Brabant}}
  - Wage{2072.0}
- ▶ xebib-dusop-begen-cyvas-lurid-tydeh-sozig-caru...
- ▶ xebon-gebig-lokiz-ramip-kokim-rycep-nyrin-ziliv-...
- ▶ xebim-lypel-dulyp-hecal-gytag-rofih-fekiz-guguh...
- ▶ xebok-gesur-gafyh-kynan-cihap-diryp-nepof-bob...
- ▶ xebof-myfun-ruvif-zemip-rypem-bypok-feduz-su...

## StatutenRegister:Data

## Records

- ▶ 50090272156
- ▶ 07111157873
- ▶ 00061717330
- ▶ 45051810755
- ▶ 93062824964
- ▼ \*31071211667
  - Name{Marion Peeters}
  - Professional Statute{Zelfstandige in Bijberoep}
  - Wage{2072.0}
- ▶ 92022424788

## StatutenRegister:AnonDataDoma...

## Records

- ▶ xebic-gyvav-dypih-vukav-donyk-nalag-huvel-kuk...
- ▶ xebib-kinyv-kahyf-balir-fuduc-hakot-hihoh-ryrop...
- ▶ xebin-dahup-fezig-vumuc-zucyf-fuvim-fazir-nobo...
- ▶ xebon-pogov-hifit-bizop-pafes-zozun-cigiv-dyfar...
- ▼ \*xebir-hirom-bozan-cegyn-valyg-zucad-cynyp-ni...
  - Professional Statute{Zelfstandige in Bijberoep}
  - Wage{2072.0}
- ▶ xebos-mamyg-dubut-kenin-fusov-todav-byren-c...
- ▶ xeboc-lulyn-sirak-fufed-syhar-pofid-dutyb-nokim...
- ▶ xebom-damov-tetuk-hvum-hvmv...

## Fod Economie: project Beta's Fin...

## Records

- ▶ xebol-girek-luper-vymav-pyzok-kebev-giuan-tese...
- ▶ xebif-musel-cezol-batin-fazys-zabuh-vopyv-helet...
- ▶ xebip-fizor-mizal-lohif-zacib-vylem-motos-hidyr-...
- ▶ xebik-mugyg-guzor-ticin-puvyv-cavek-kymub-pu...
- ▶ xebos-zyvyb-kuber-kyrus-suzuk-bavep-conyl-bit...
- ▶ xebid-kazav-biheb-gumob-vacip-lyhud-zotel-vac...
- ▼ \*xebit-kyrom-fozez-rizyd-tyfog-getam-lasol-pulin...
  - BirthDate{1931}
  - Domicile{3470;Ransberg}

# Proof-of-Concept

Het theoretisch model werkt ook in de praktijk

## Performance pseudonym conversion

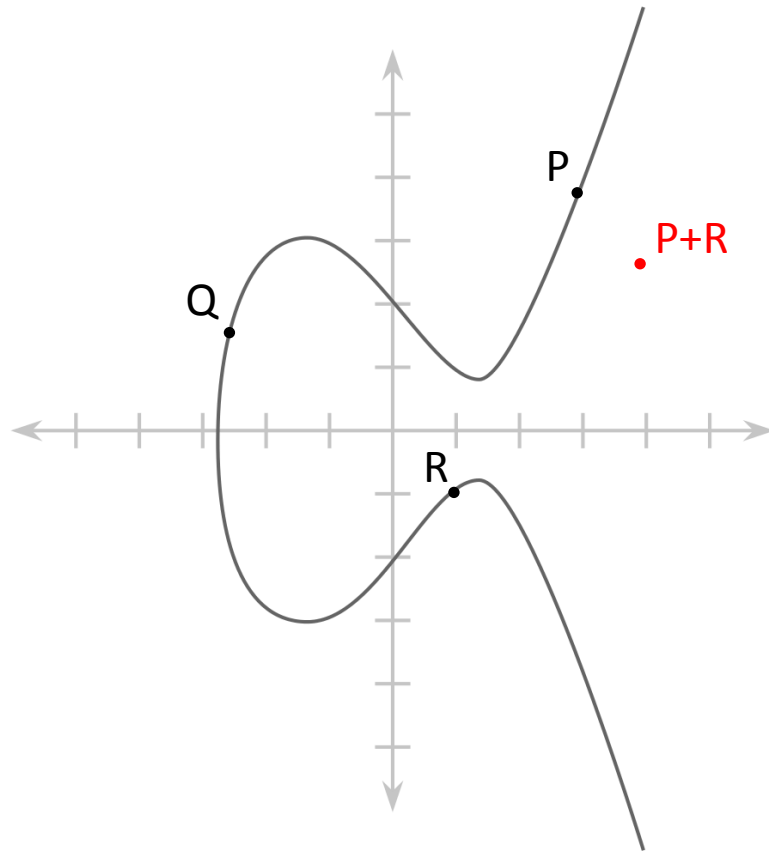
*PC Windows 7 Enterprise (64bit) op één 2,66Ghz Intel i5 core*

RSA			EC		
Key size	One operation	Ops / hour	Key size	One operation	Ops / hour
1536 bit	58ms	62070	192 bit	0,4ms	9 million
2048 bit	135ms	26700	224 bit	0,6ms	6 million
3072 bit	440ms	8180	256 bit	0,7-0,8ms	4-5 million

# Een beetje cryptografie



# Elliptische Krommen (EC)



$$P = (x_P, y_P)$$

$$Q = (x_Q, y_Q)$$

$$R = (x_R, y_R)$$

## De curve is een groep

We willen een bewerking (+)  
zodat telkens  $P + R \in \text{curve}$

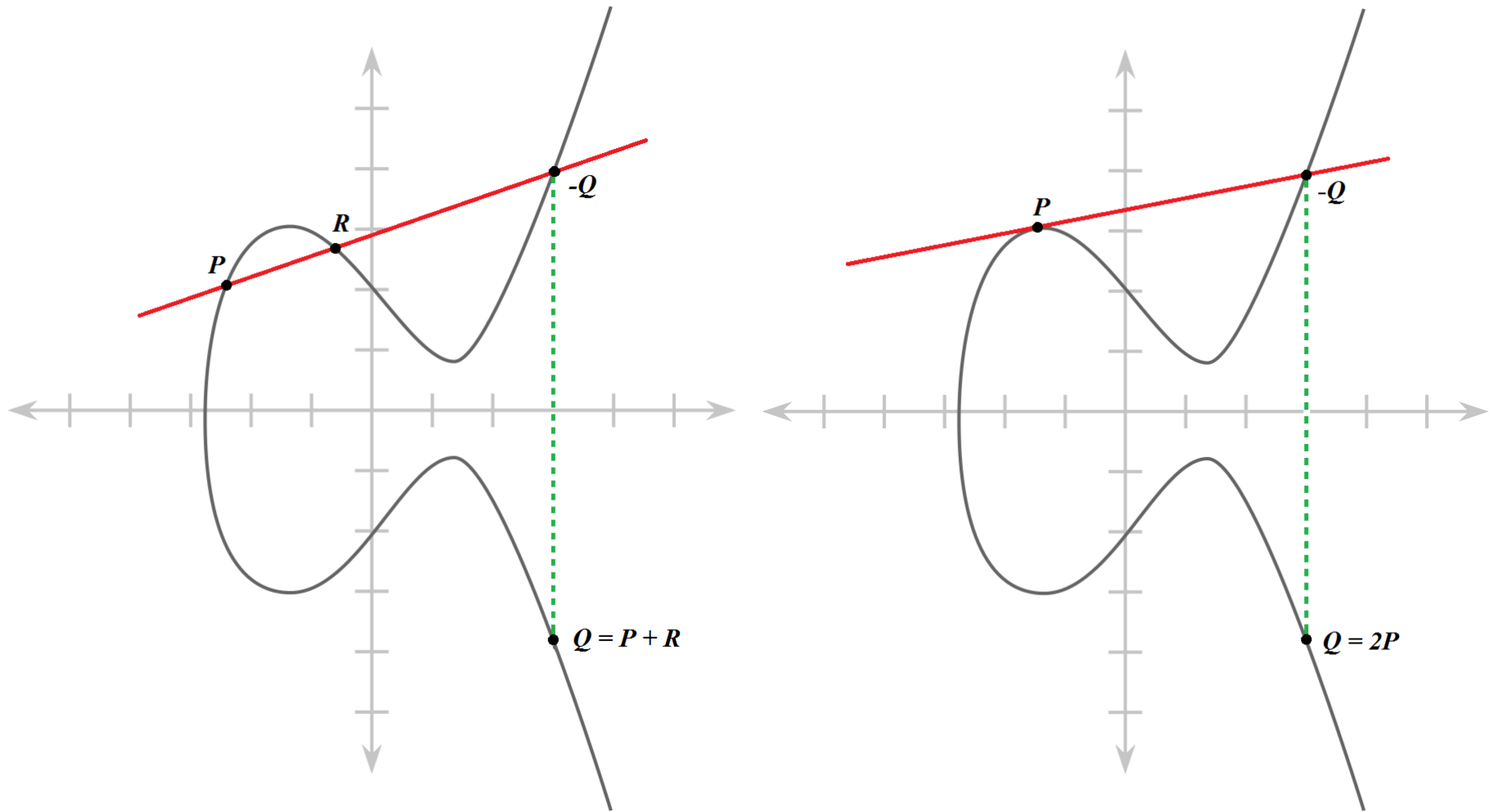
Wat werkt niet:

$$P + R = (x_P + x_R, y_P + y_R)$$

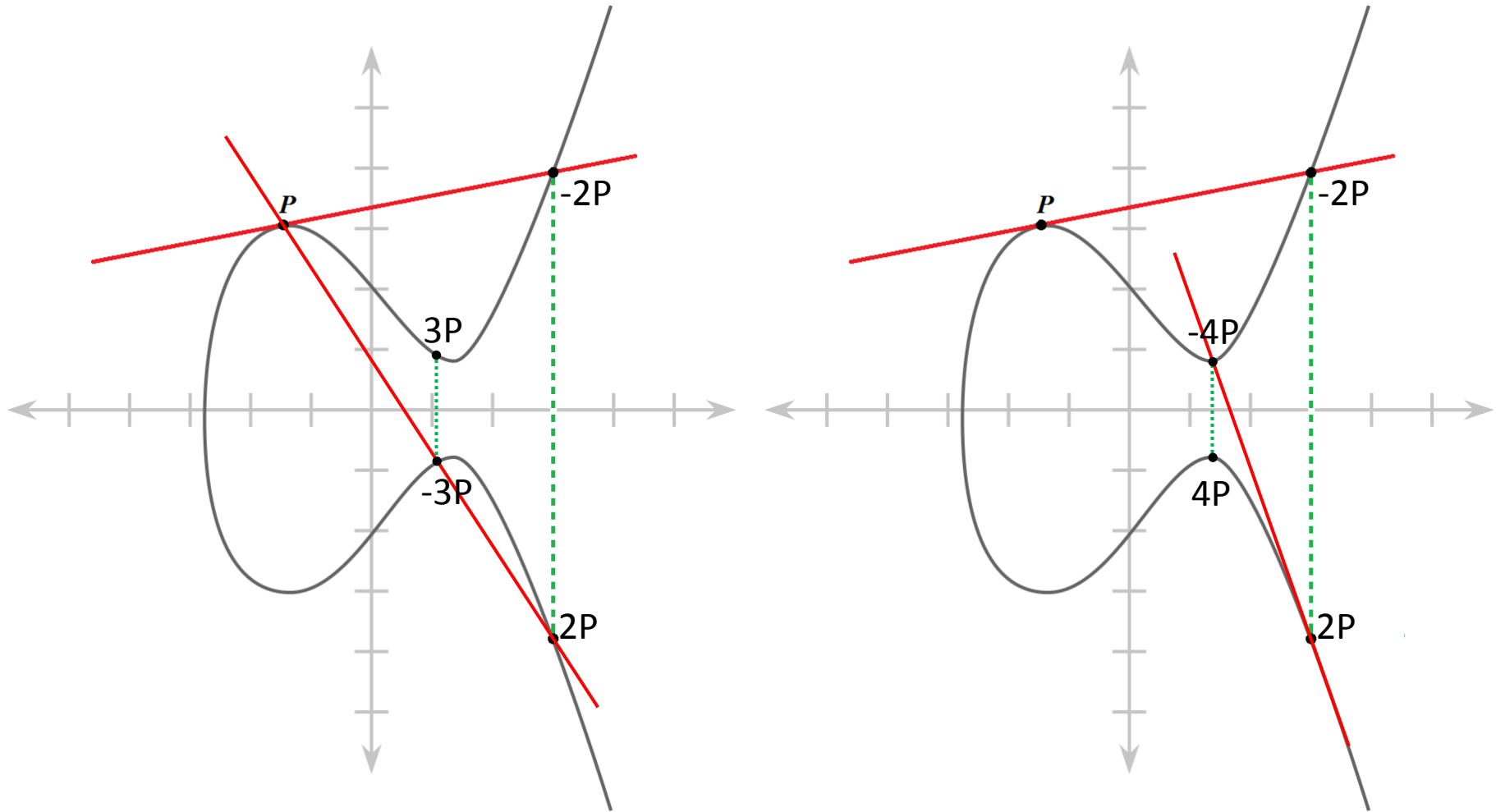
## Geschiedenis

- Voorgesteld in 1985 door Koblitz & Miller
- Wijdverspreid sinds 2004 – 2005
- Graduele shift van RSA naar EC (want efficiënter)

# Elliptische Krommen (EC)



# Elliptische Krommen (EC)



**Easy:**  $Q \leftarrow n \cdot P = \underbrace{P + P + \dots + P}_{n \text{ times}} \quad n \in \mathbb{Z}$

**Hard:**  $n \leftarrow P, Q$

# Central Idea



$$DP \leftarrow s_d \cdot I$$



$$TP \leftarrow \frac{s_t}{s_d} \cdot DP = s_t \cdot I$$



$$PP \leftarrow \frac{s_p}{s_t} \cdot TP = s_p \cdot I$$



$$DP' \leftarrow s'_d \cdot I$$



$$TP' \leftarrow \frac{s'_t}{s'_d} \cdot DP' = s'_t \cdot I$$



$$PP \leftarrow \frac{s_p}{s'_t} \cdot TP' = s_p \cdot I$$

**Easy:**  $Q \leftarrow n \cdot P = \underbrace{P + P + \dots + P}_{n \text{ times}} \quad n \in \mathbb{Z}$

**Hard:**  $n \leftarrow P, Q \quad n \text{ times}$

# Deanonimisatie

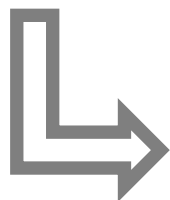


# Fraudebestrijding

Onderzoeker moet in staat zijn verdacht gedrag te detecteren

Onderzoeker mag betrokken burger niet kunnen identificeren

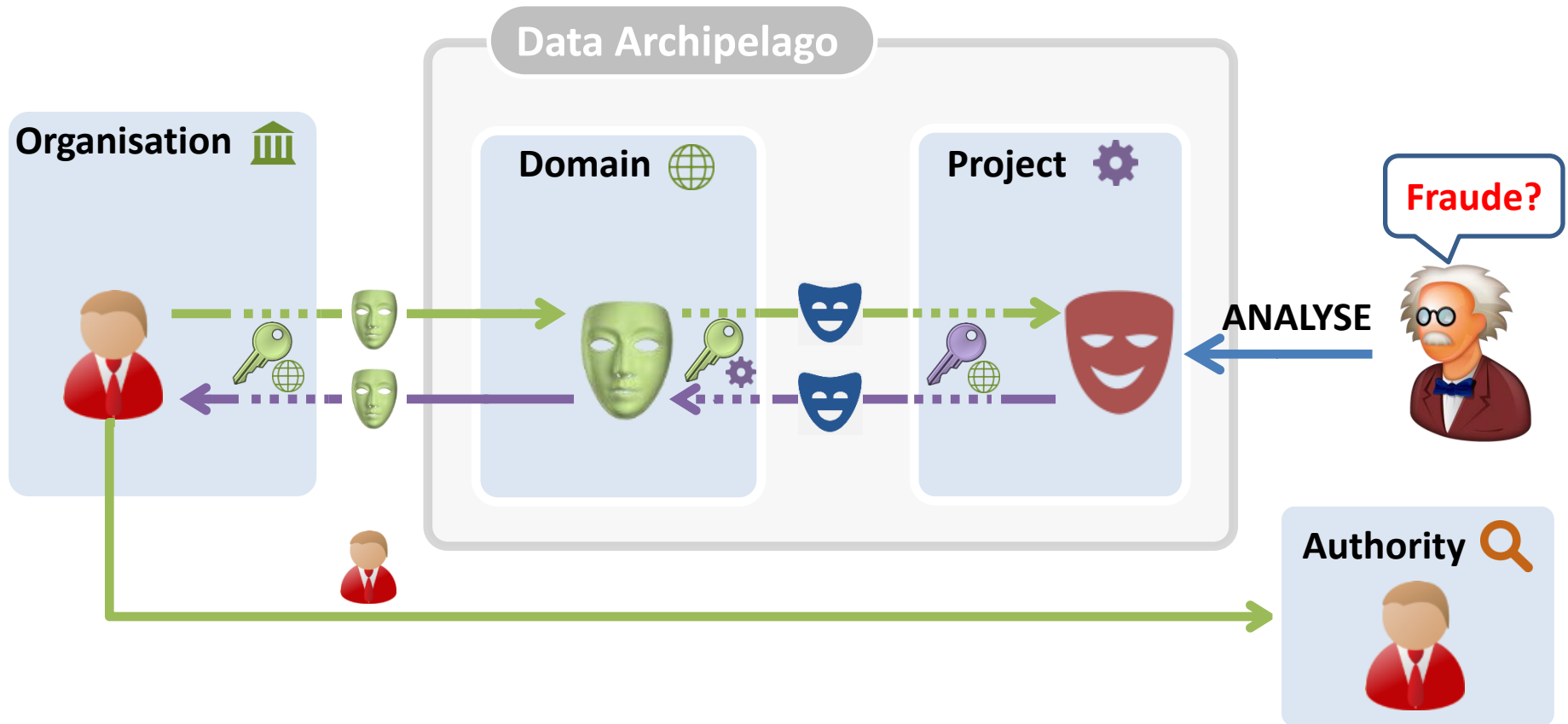
Enkel een specifieke autoriteit mag een pseudoniem kunnen linken aan burger



**Geïntegreerde & flexibele oplossing**

# Org. deanonimiseert

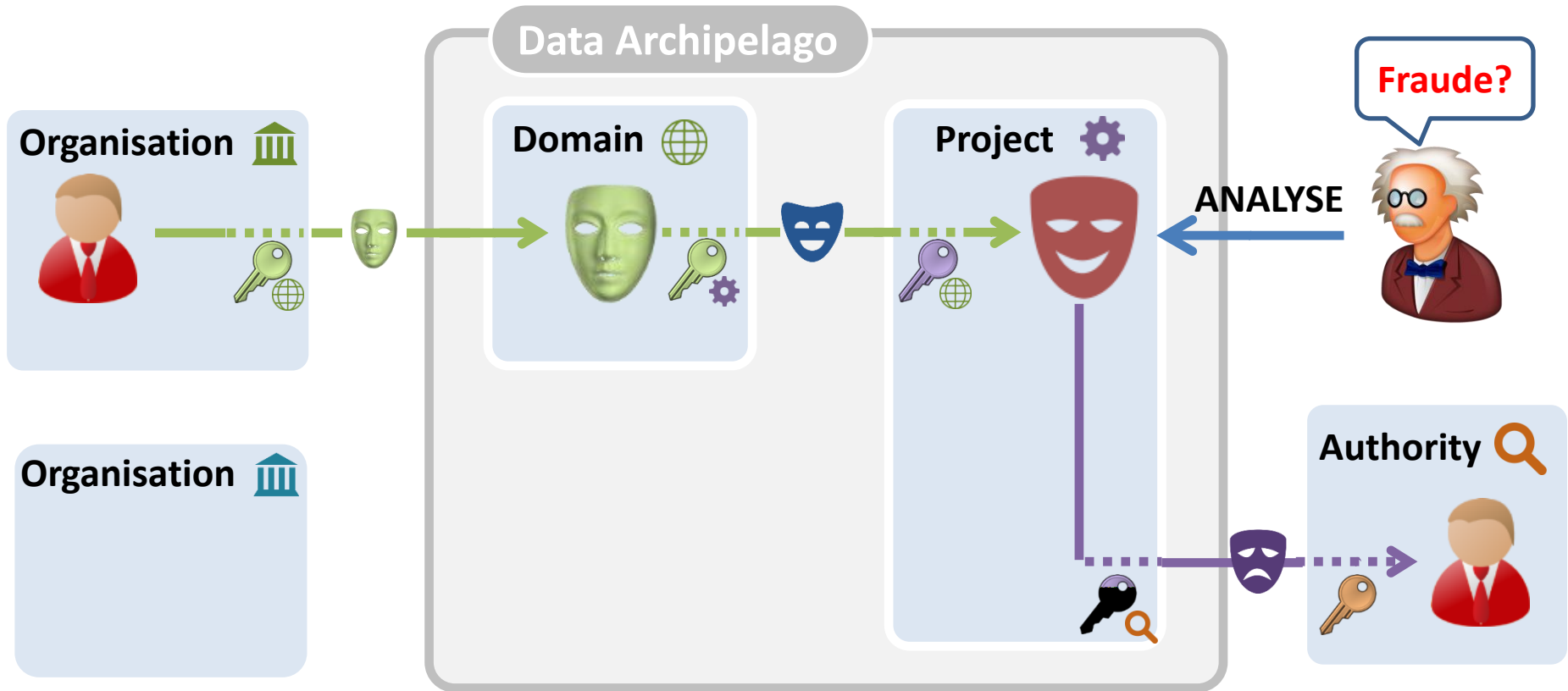
1



Betrokken organisatie leert de identiteit van de burger en neemt verdere stappen

# Geen org. betrokken

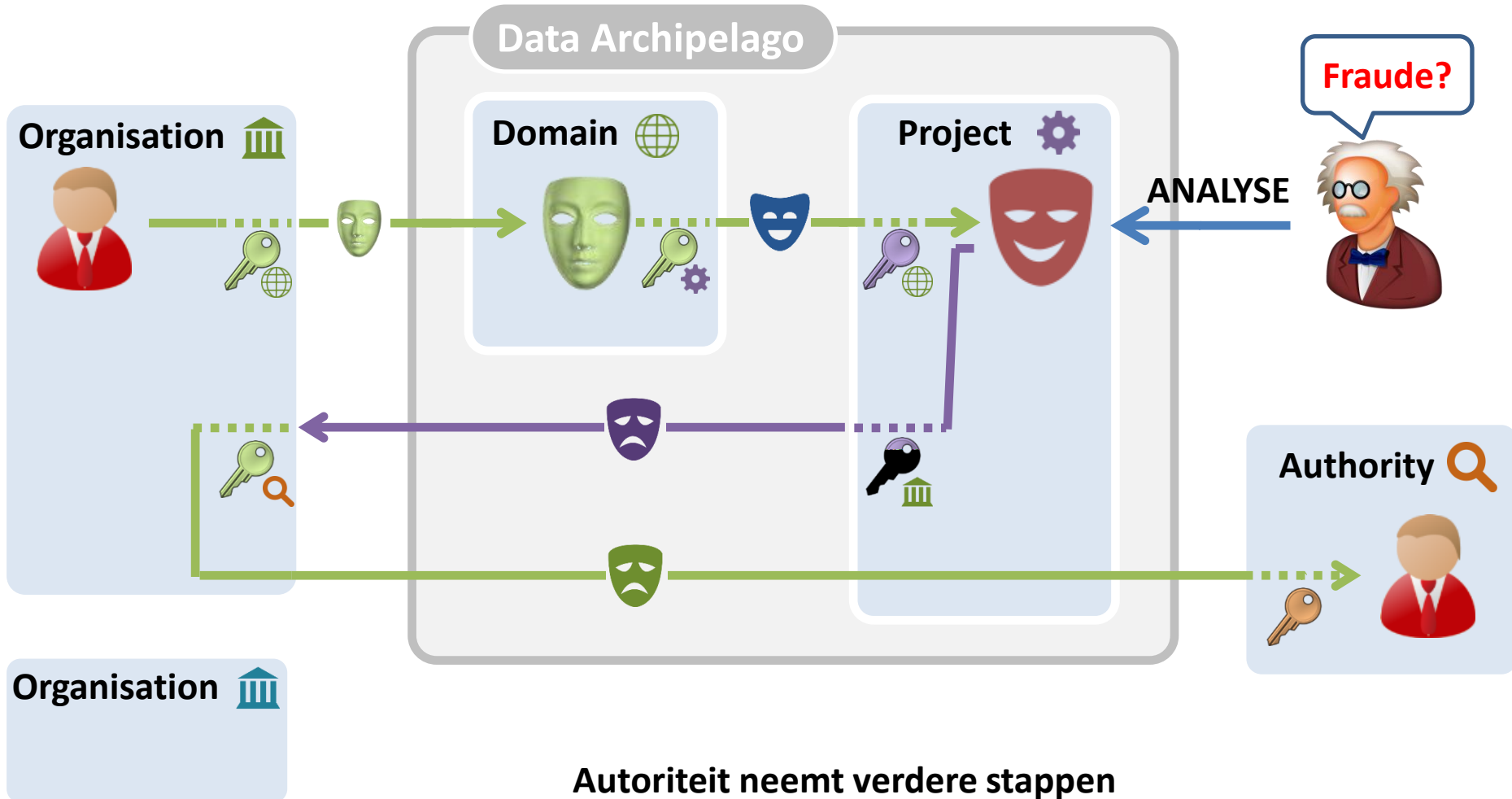
2



Deanonimisatie gebeurt zonder medeweten van betrokken organisaties  
Autoriteit neemt verdere stappen

# Eén org. keurt goed, maar leert identiteit burger niet

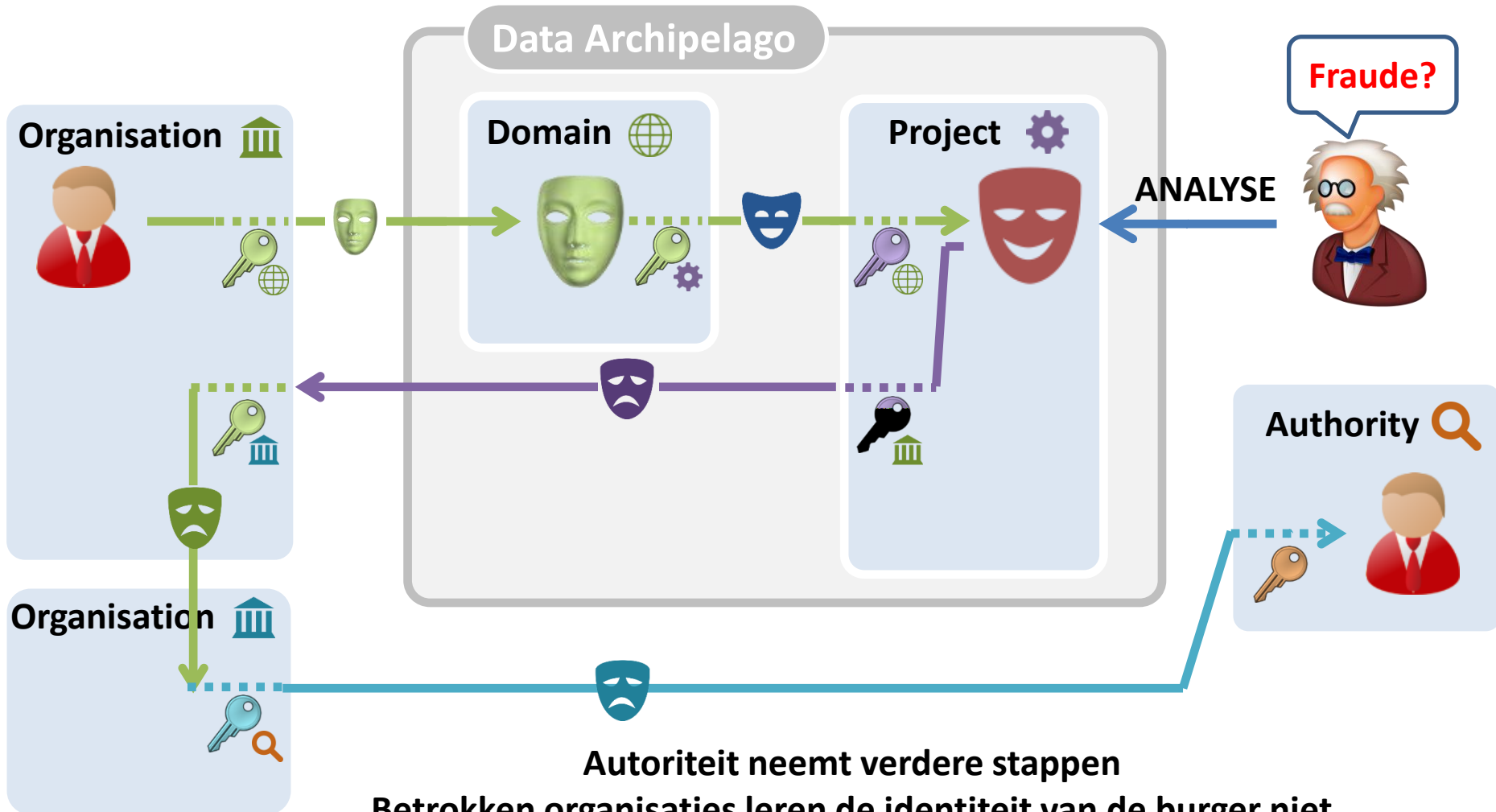
3



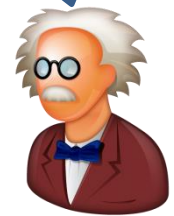
Autoriteit neemt verdere stappen  
Betrokken organisaties leren de identiteit betrokken burger niet  
Eén betrokken organisatie keurt de deanonimisatie goed

# Meerdere orgs. keuren goed zonder identiteit te leren

4

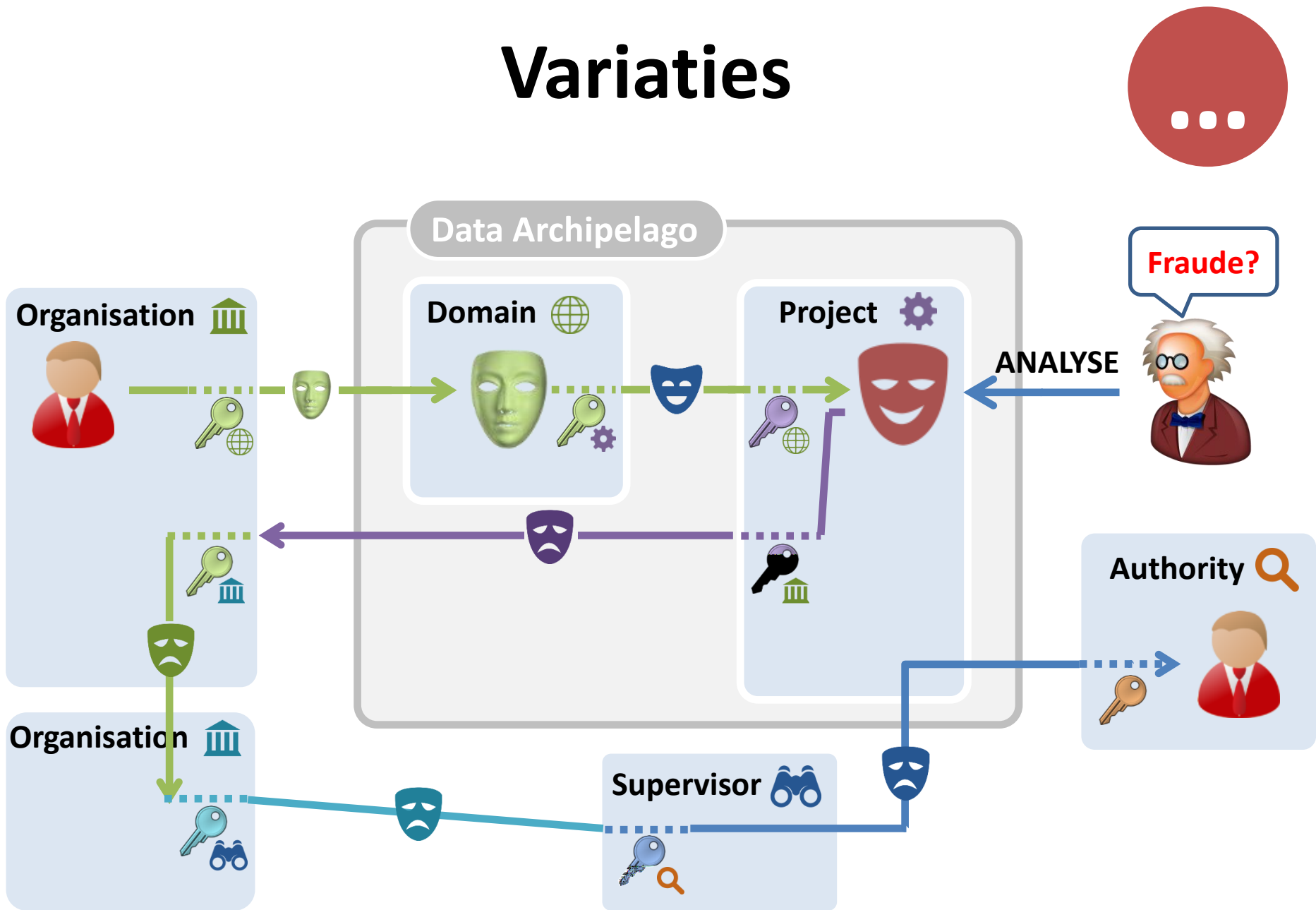


Fraude?



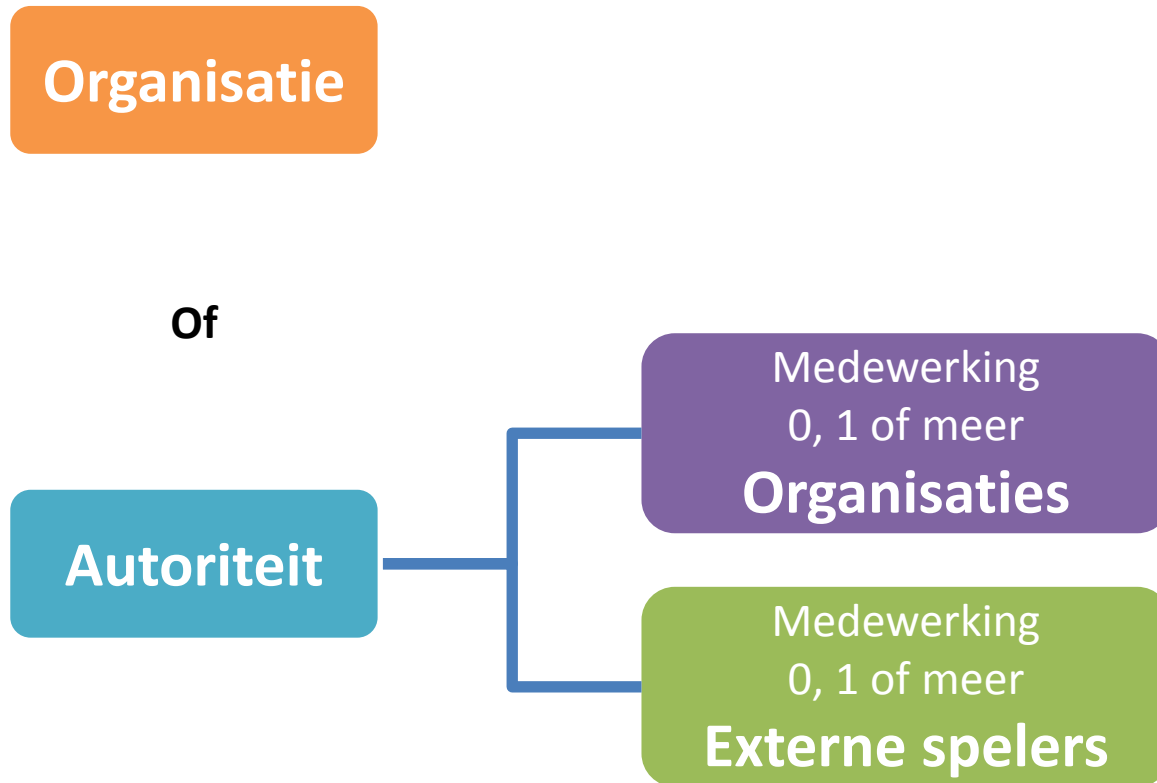
**Autoriteit neemt verdere stappen**  
**Betrokken organisaties leren de identiteit van de burger niet**  
**Meerdere betrokken organisaties keuren deanonimisatie goed**

# Varieties



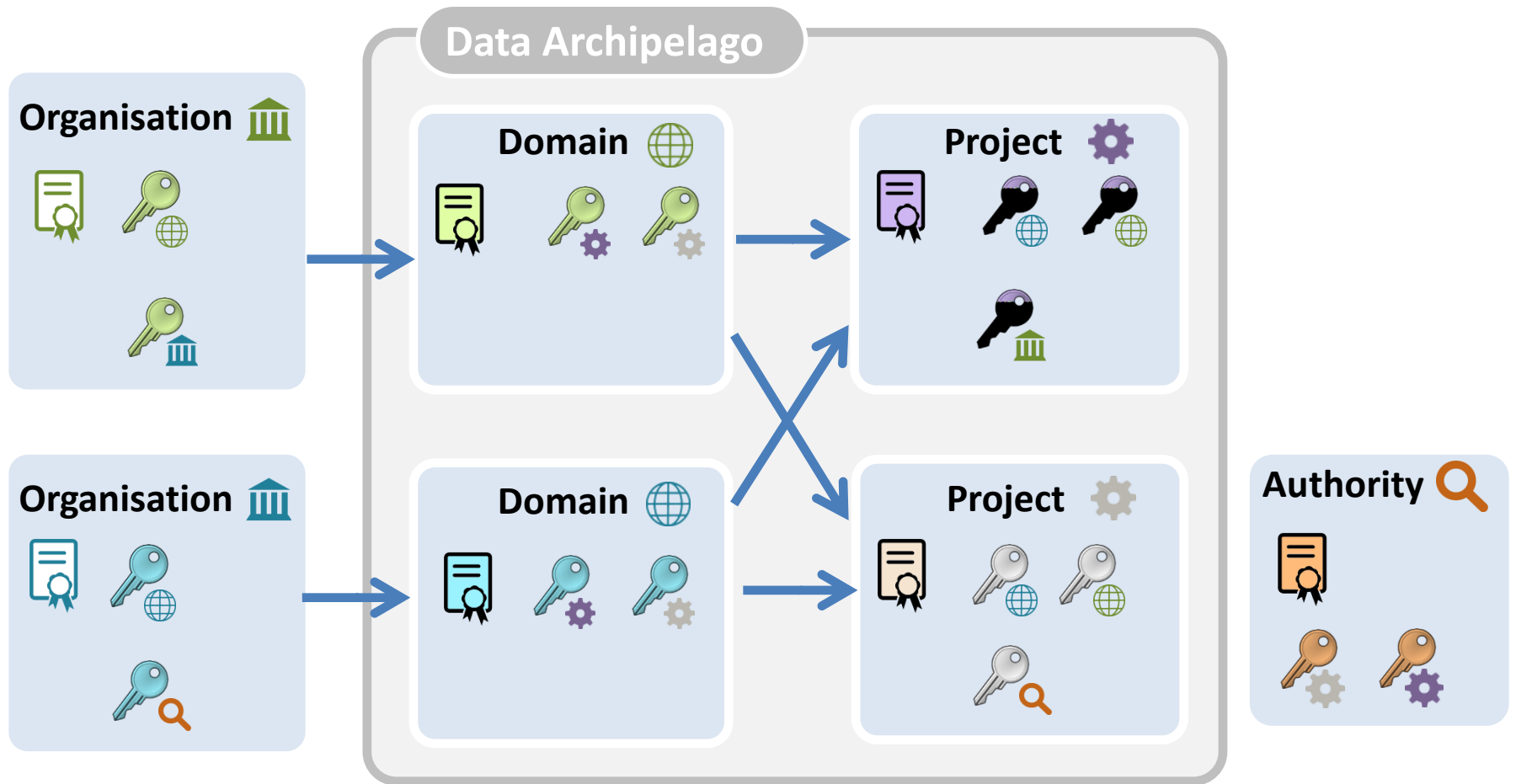
# Samenvatting

## Flexibele & efficiënte deanonimisatie

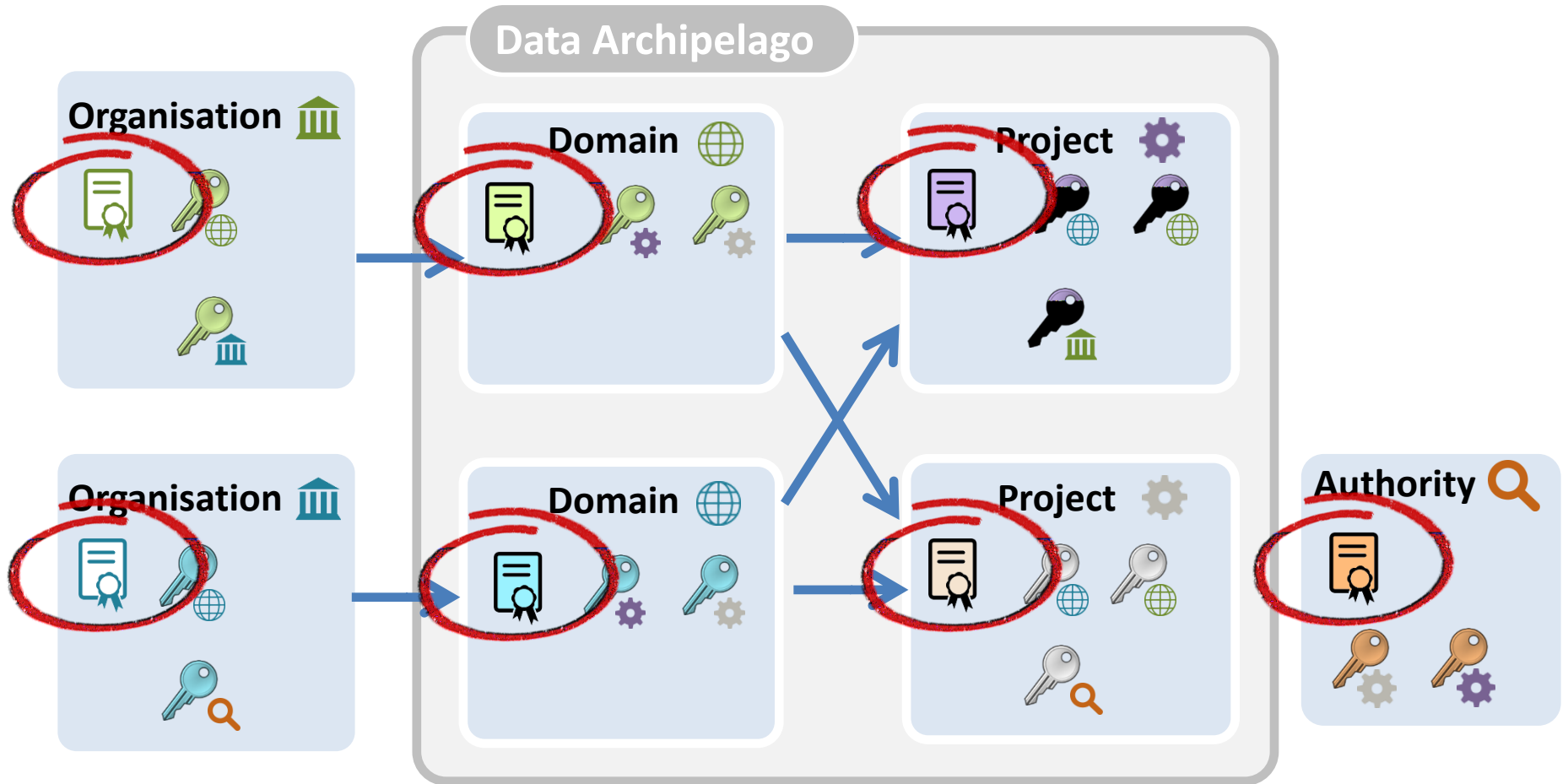


# Sleutelbeheer

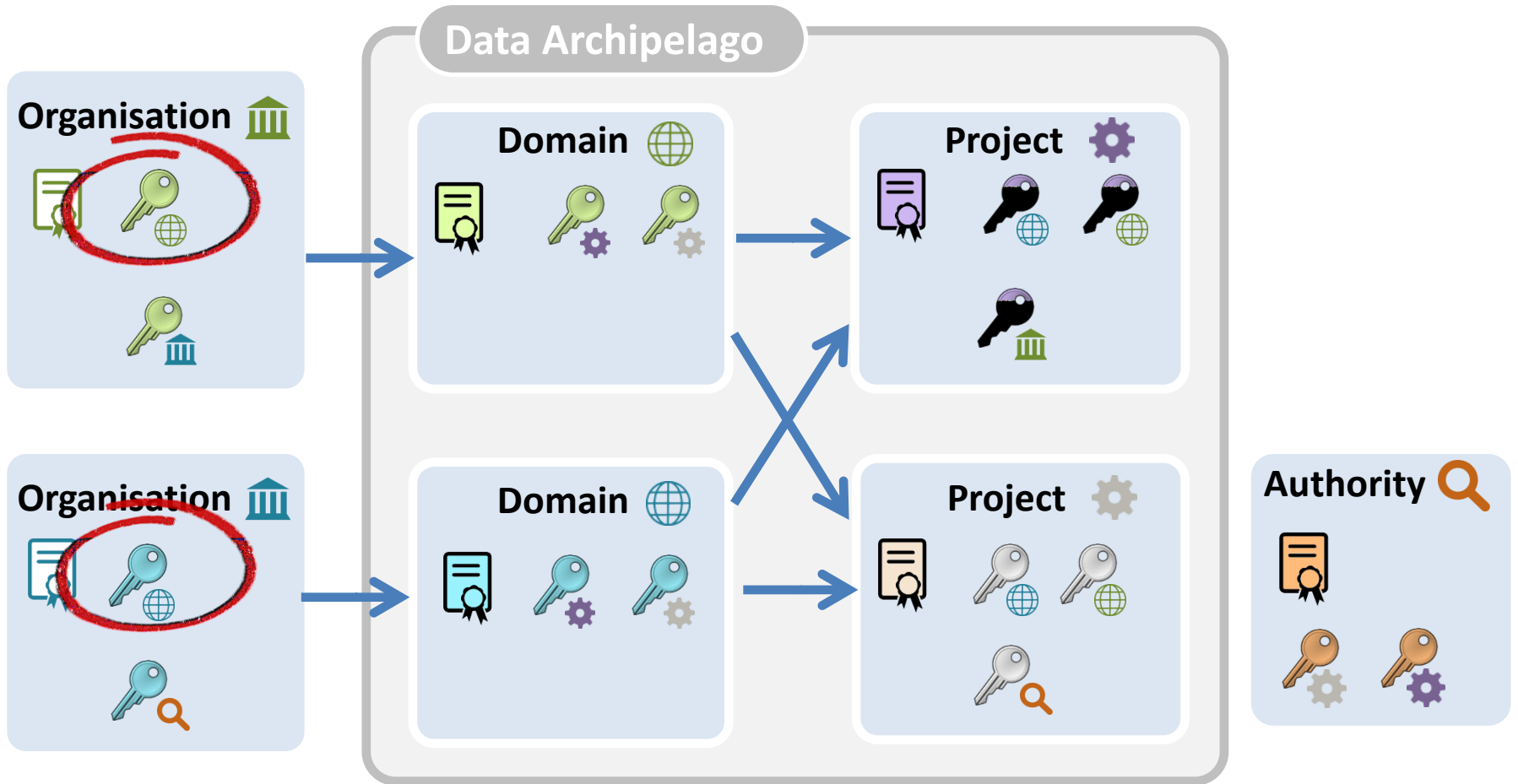
# Key Properties



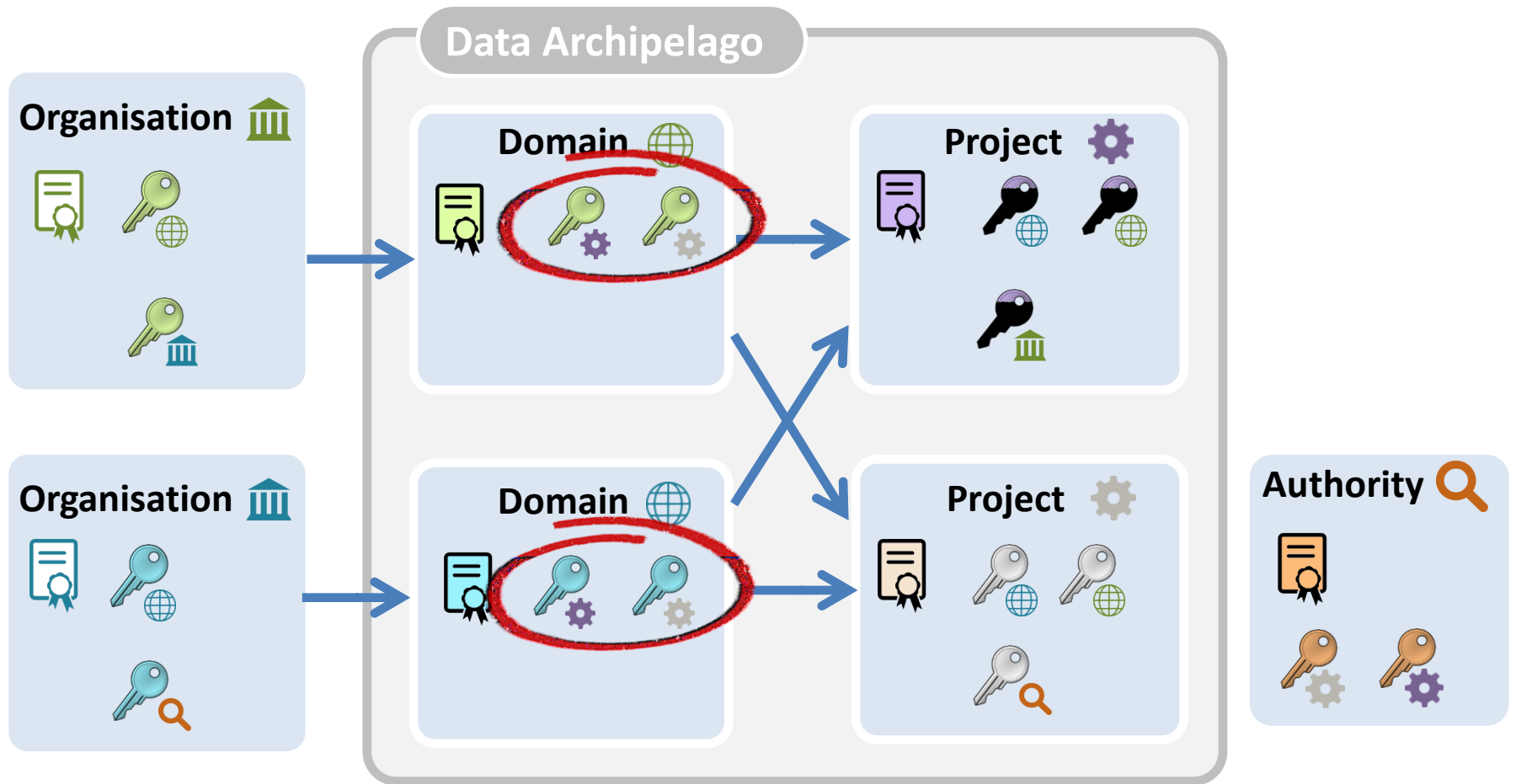
# Sleutels



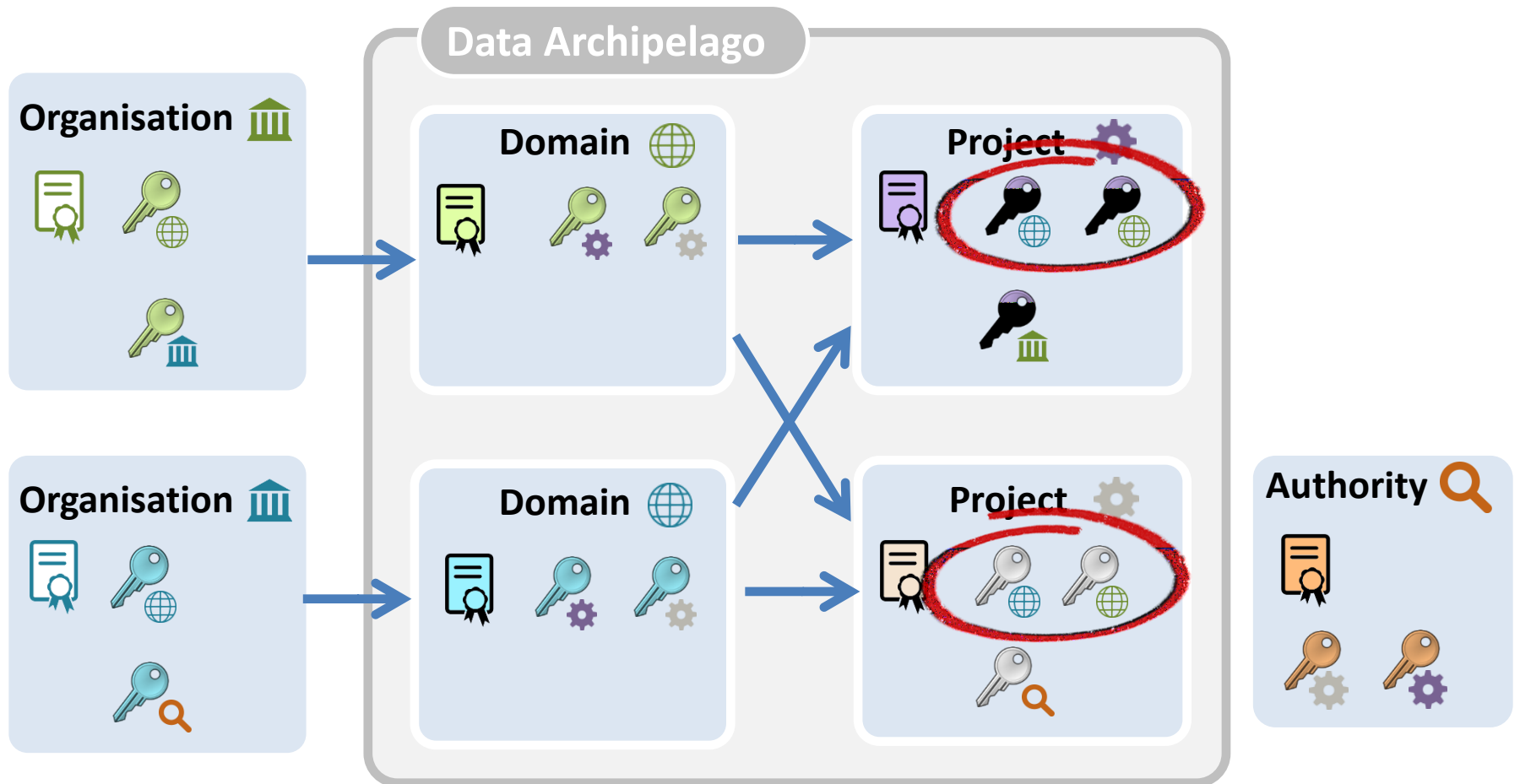
# Sleutels



# Sleutels

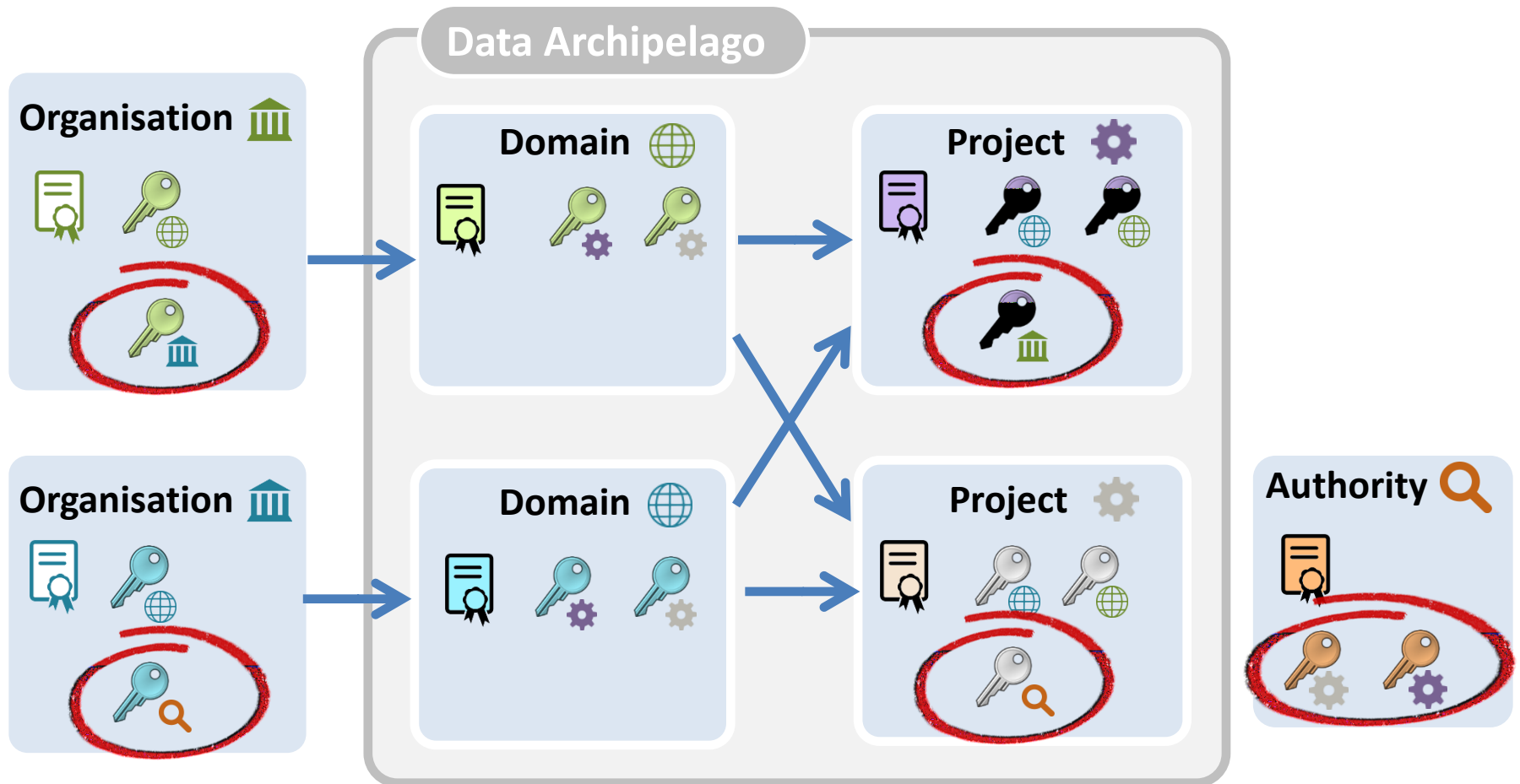


# Sleutels



Kan soms al snel verwijderd worden

# Sleutels



Opslag sleutels?

Sleutellengte?

Sleutelgeneratie?

# Opslag sleutels

## Software-based

- V.b. PKCS#12
- Password given to very slow function

## Hardware-based (HSM)

Network attached



Server embedded



Portable / USB



Cryptografische operaties in HSM → sleutel verlaat HSM nooit

# Aanbevolen sleutellengte

Verschillende aanbevelingen, waaronder:

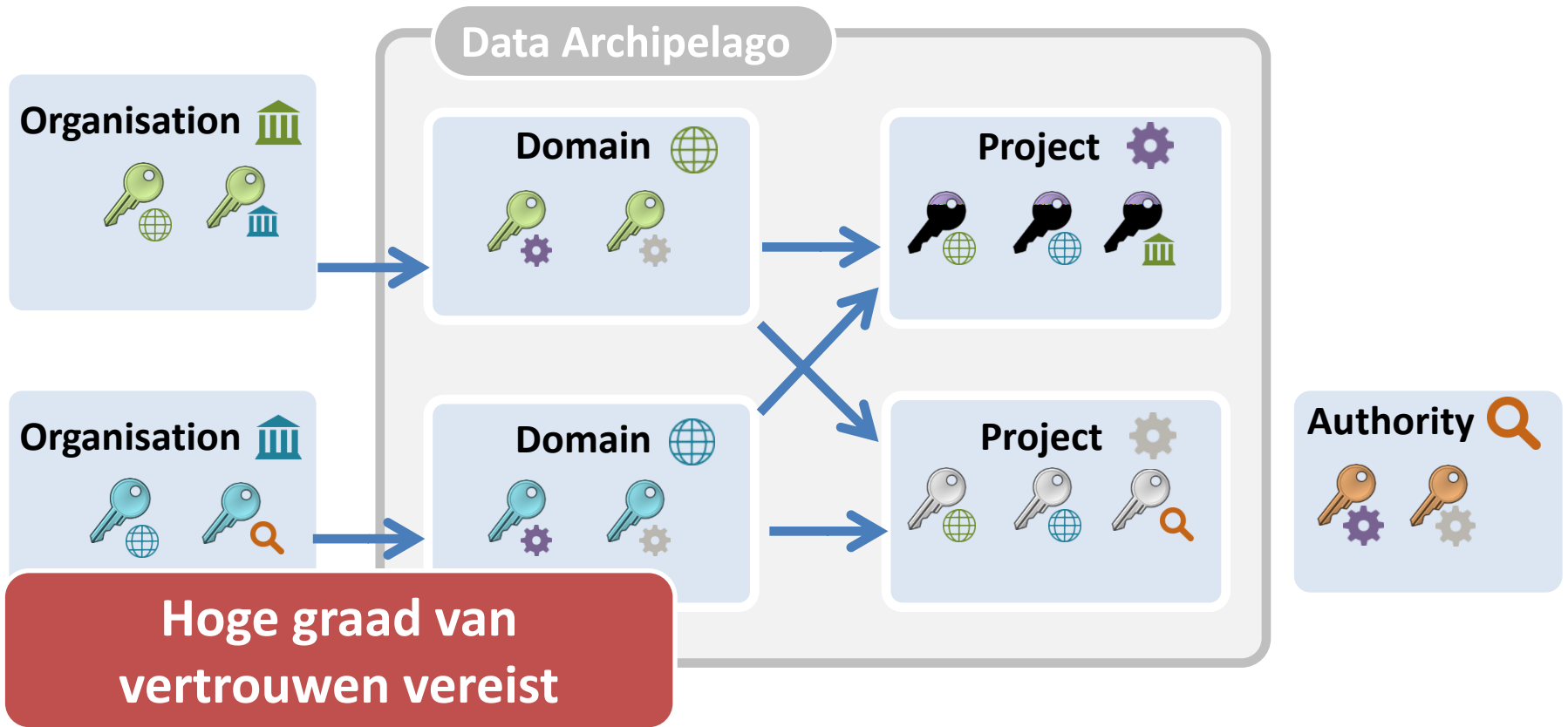
## NIST

Date	AES	RSA	EC
2010	80	1024	160
2011 - 2030	112	2048	224
> 2030	128	3072	256
>>> 2030	256	15360	512

## Lenstra

Date	AES	RSA	EC
2018	80	1329	160
2042	96	2124	192
2066	112	3154	224
2090	128	4440	256
2282	256	26268	512

# Sleutelgeneratie 1

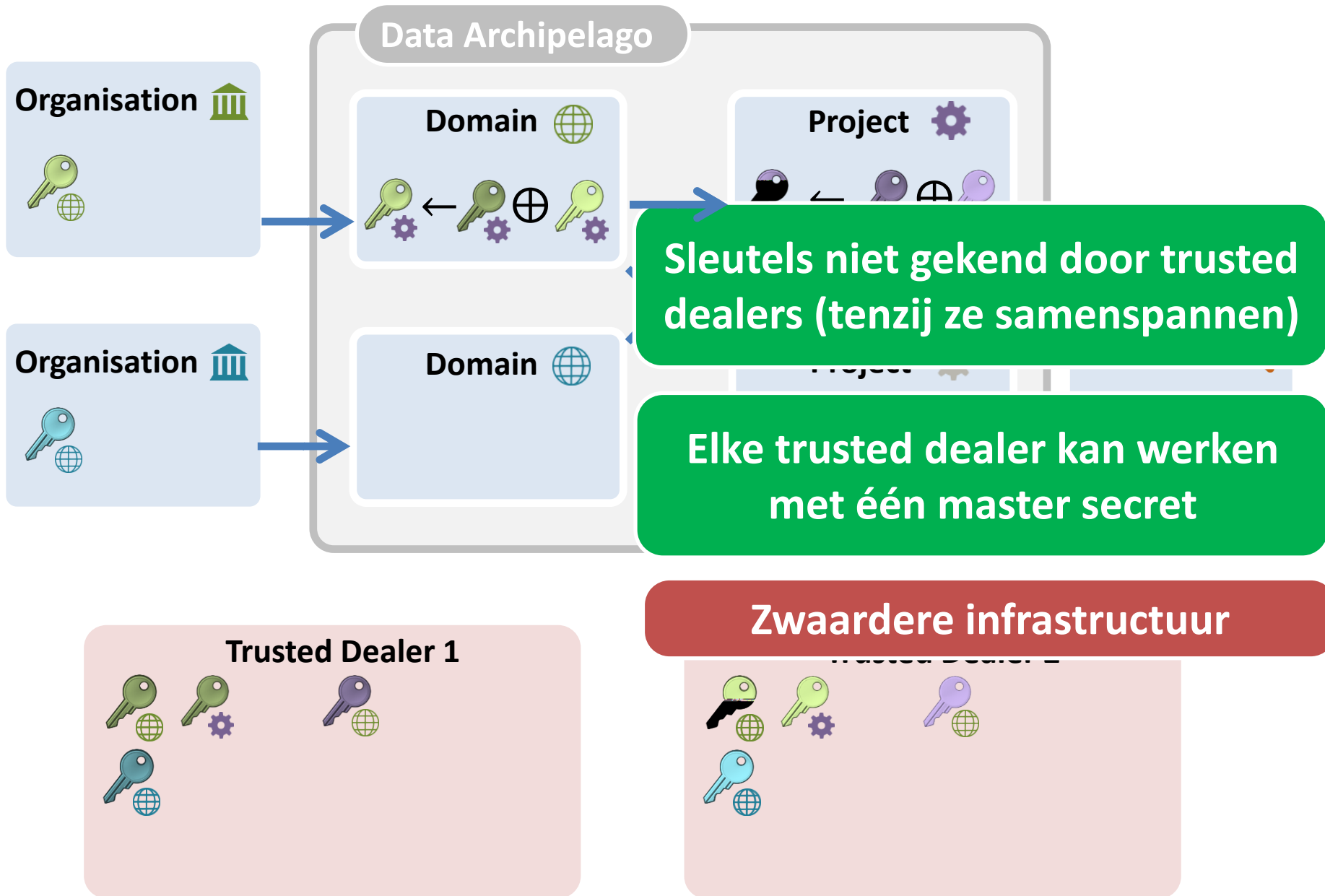


Centrale trusted dealer kent alle sleutels

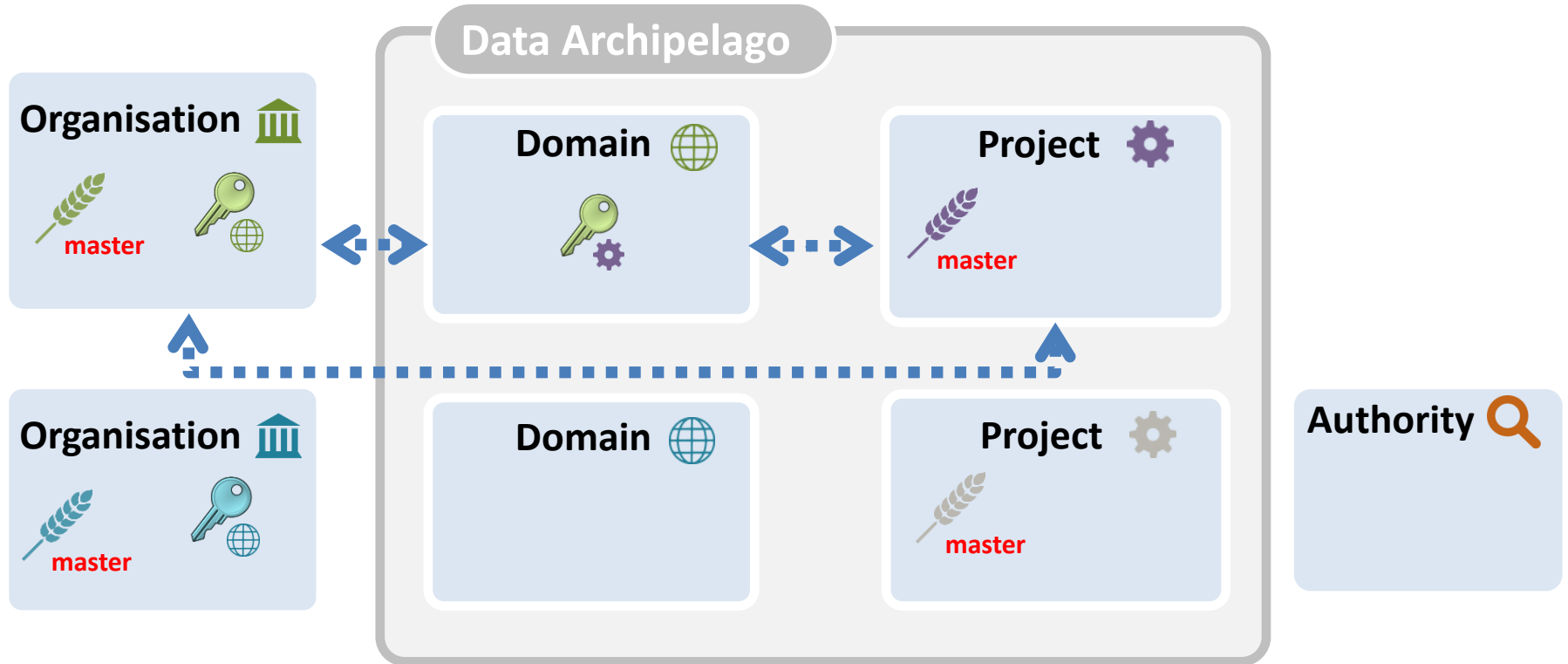
Alle sleutels afleidbaar uit één enkel master secret



# Sleutelgeneratie 2

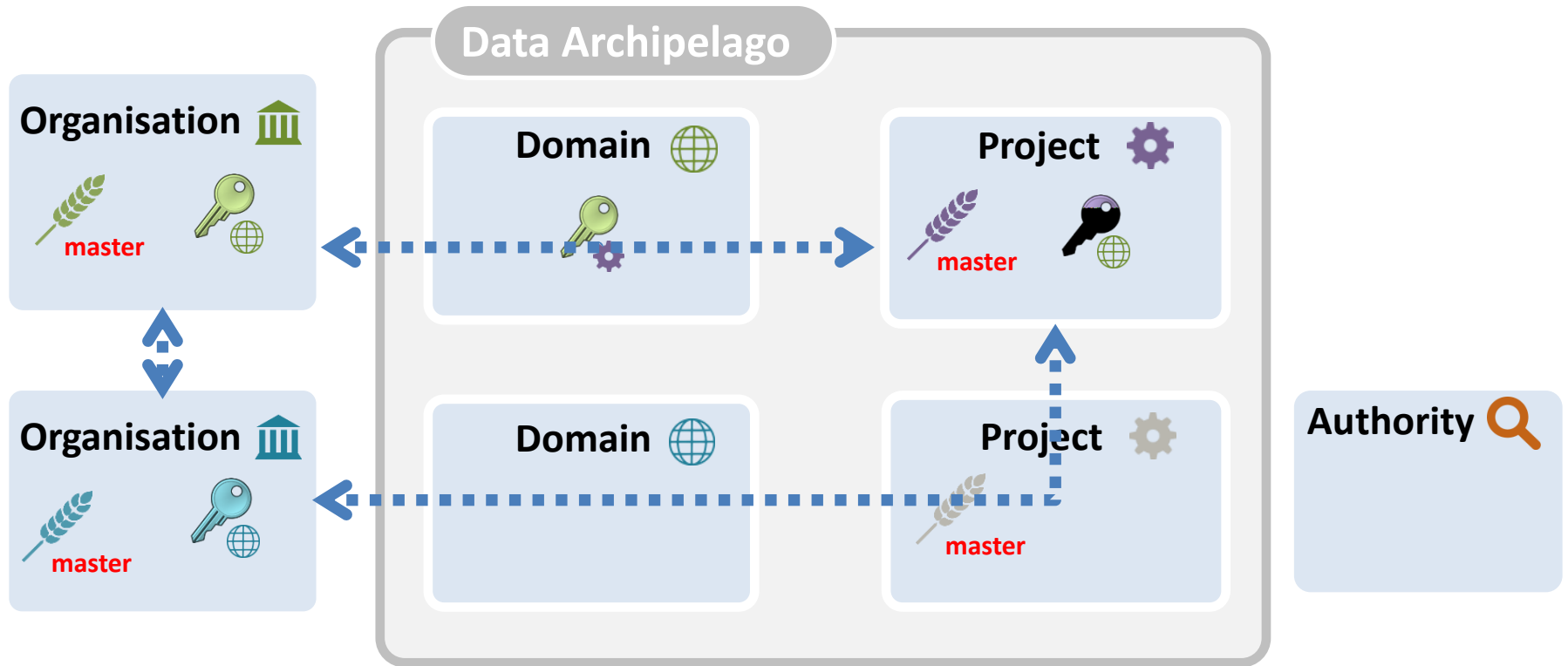


# Sleutelgeneratie 3



Domein leert niets buiten nieuwe sleutel  
Geen informatie gelekt over master secrets

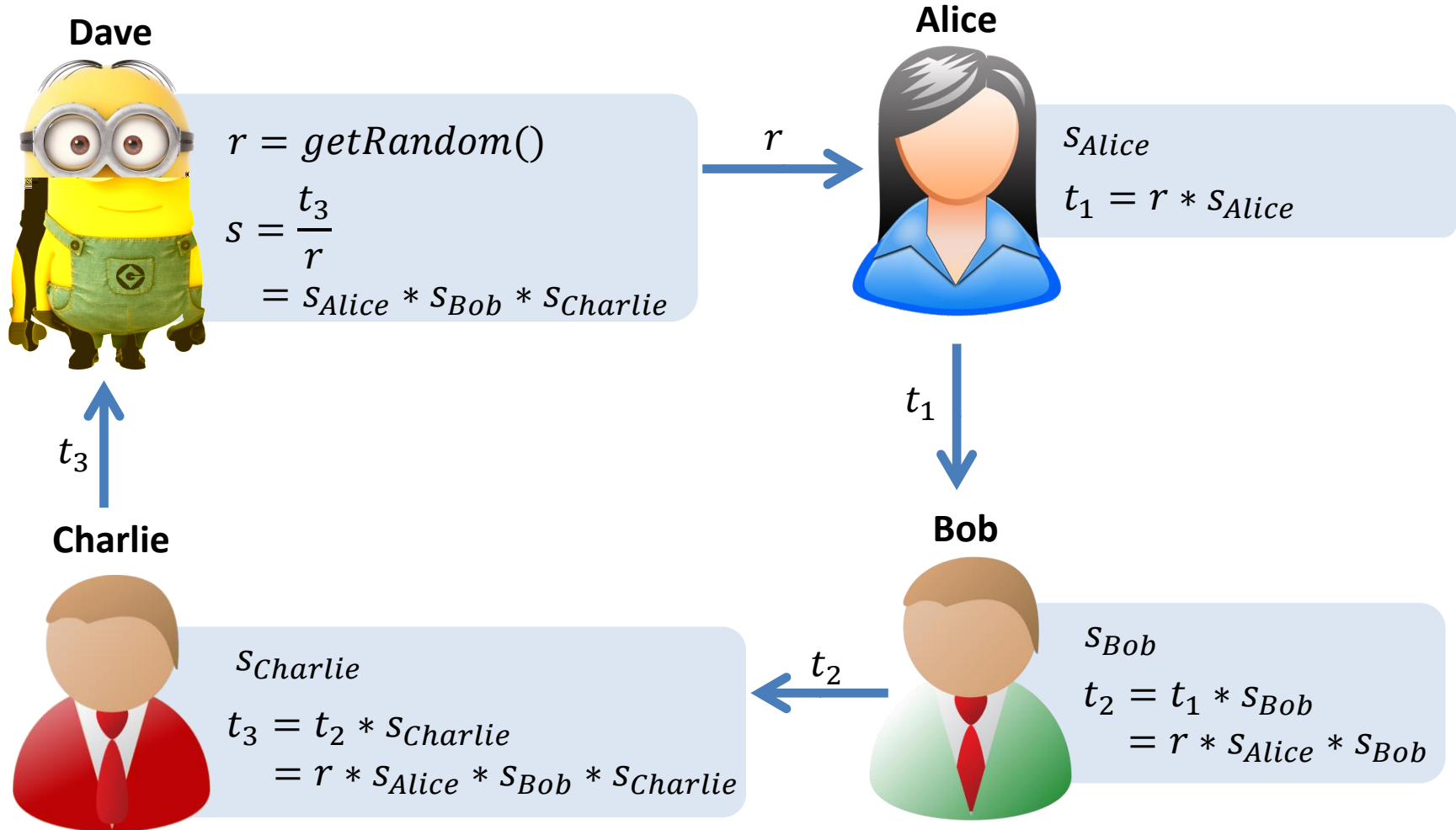
# Sleutelgeneratie 3



Project leert niets buiten nieuwe sleutel  
Geen informatie gelekt over master secrets

# Multi-Party Multiplication

Sterk vereenvoudigd!



**Dave leert  $s = S_{Alice} * S_{Bob} * S_{Charlie}$  maar niets anders**

# Generatie sleutels - Vergelijking

	Eén trusted dealer	Meerdere trusted dealers	Geen trusted dealer
<i>Vertrouwen per entiteit</i>	<b>Hoog</b>	<b>Medium</b>	<b>Laag</b>
<i>Extra partijen</i>	<b>1</b>	<b>&gt;1</b>	<b>0</b>
<i>Impact lekken geheim</i>	<b>Hoog</b>	<b>Medium</b>	<b>Laag</b>
<i>Kans lekken geheim</i>	<b>Laag</b>	<b>Medium</b>	<b>Hoog</b>
<i>Sleutel recupereerbaar</i>	<b>Ja</b>	<b>Ja</b>	<b>Ja</b>

Flexibele sleutelgeneratie en -distributie

# Samenvatting

Elke entiteit slaat  
sleutels op

Levensduur  
sleutels varieert

Veilige opslag  
vereist

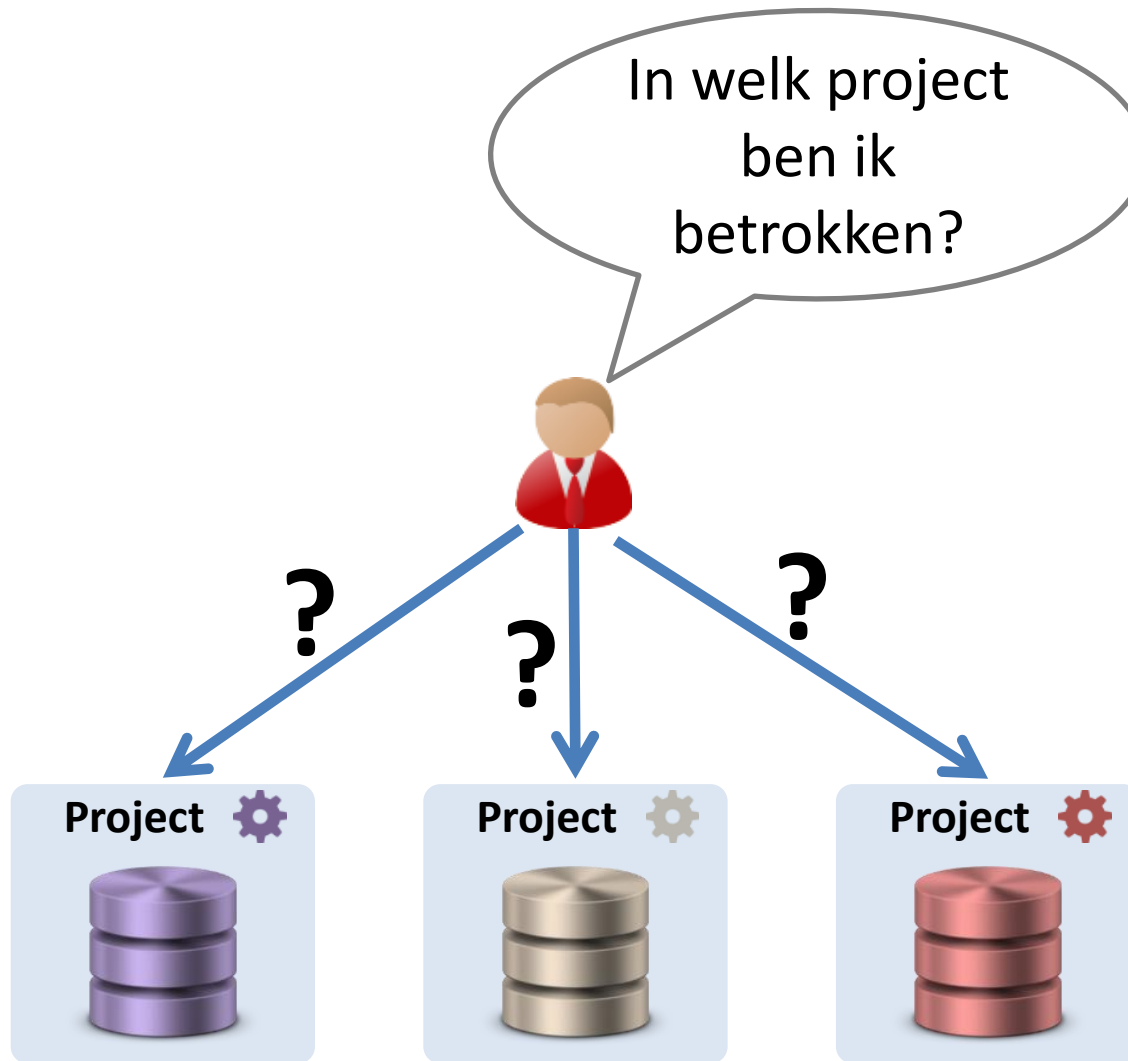
Voldoende lange  
sleutels

Flexibele sleutelgeneratie

# Transparantie



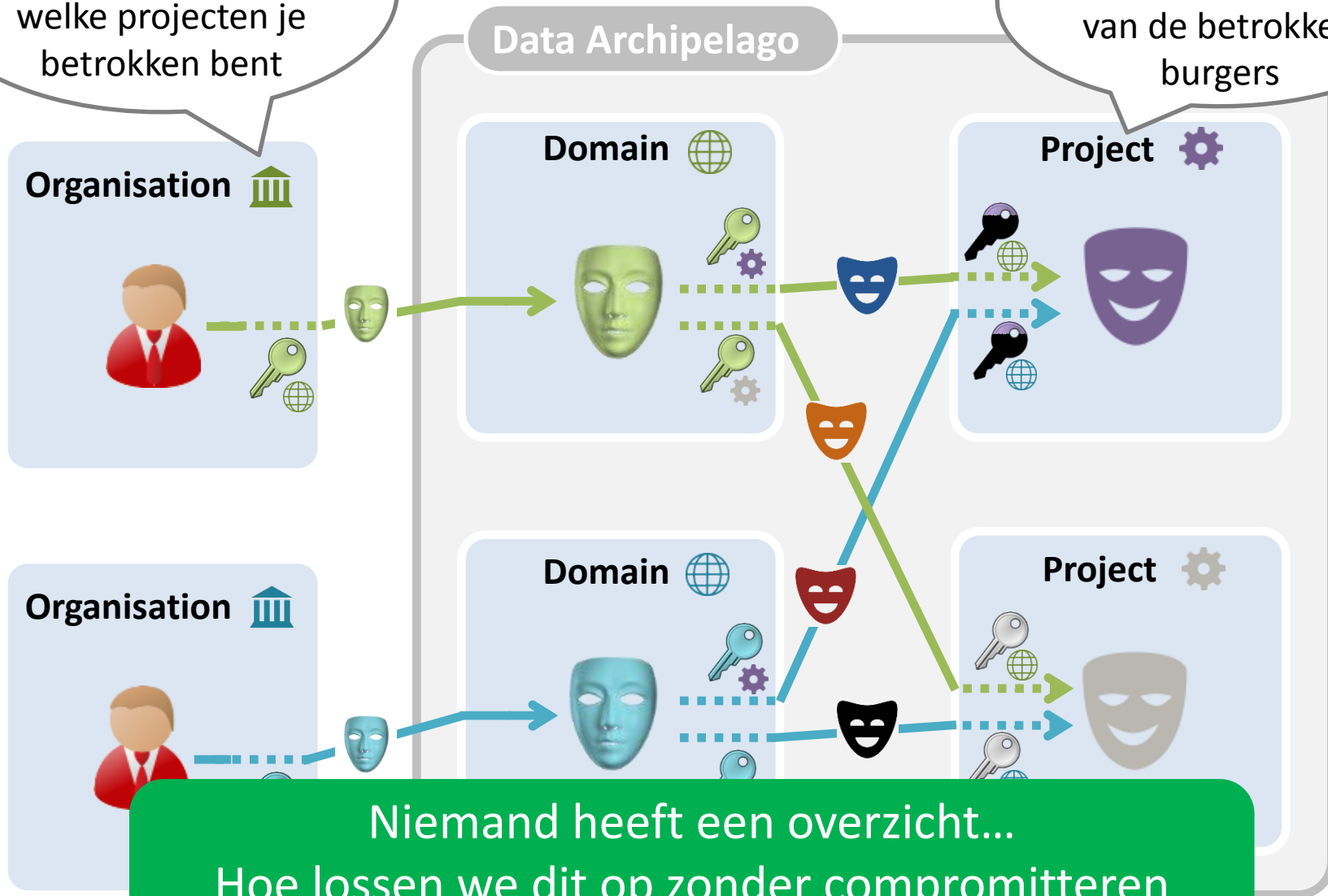
# Transparantie voor de burger



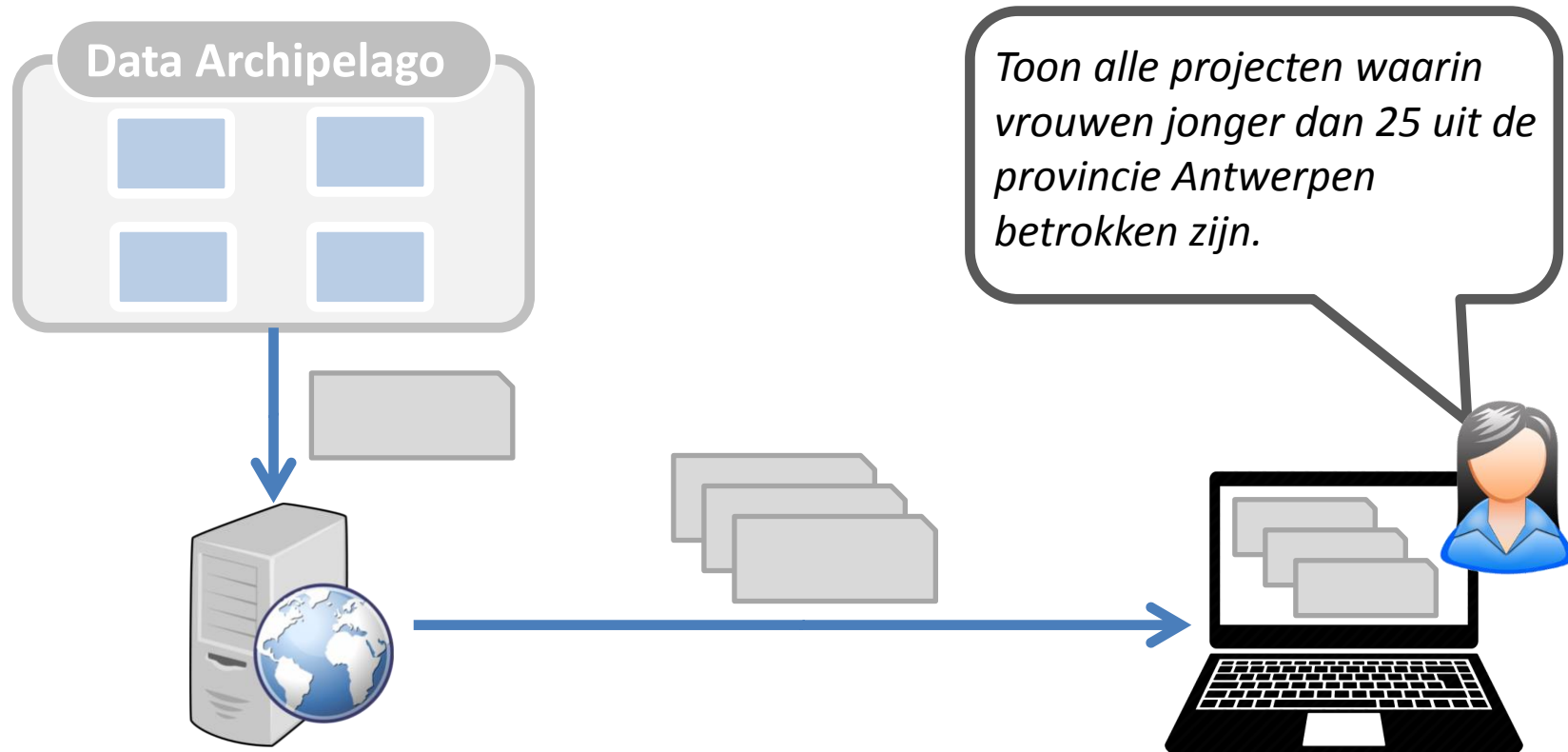
# Garantatie voor de burger

We kennen je, maar weten niet in welke projecten je betrokken bent

We kennen de identifiers niet van de betrokken burgers



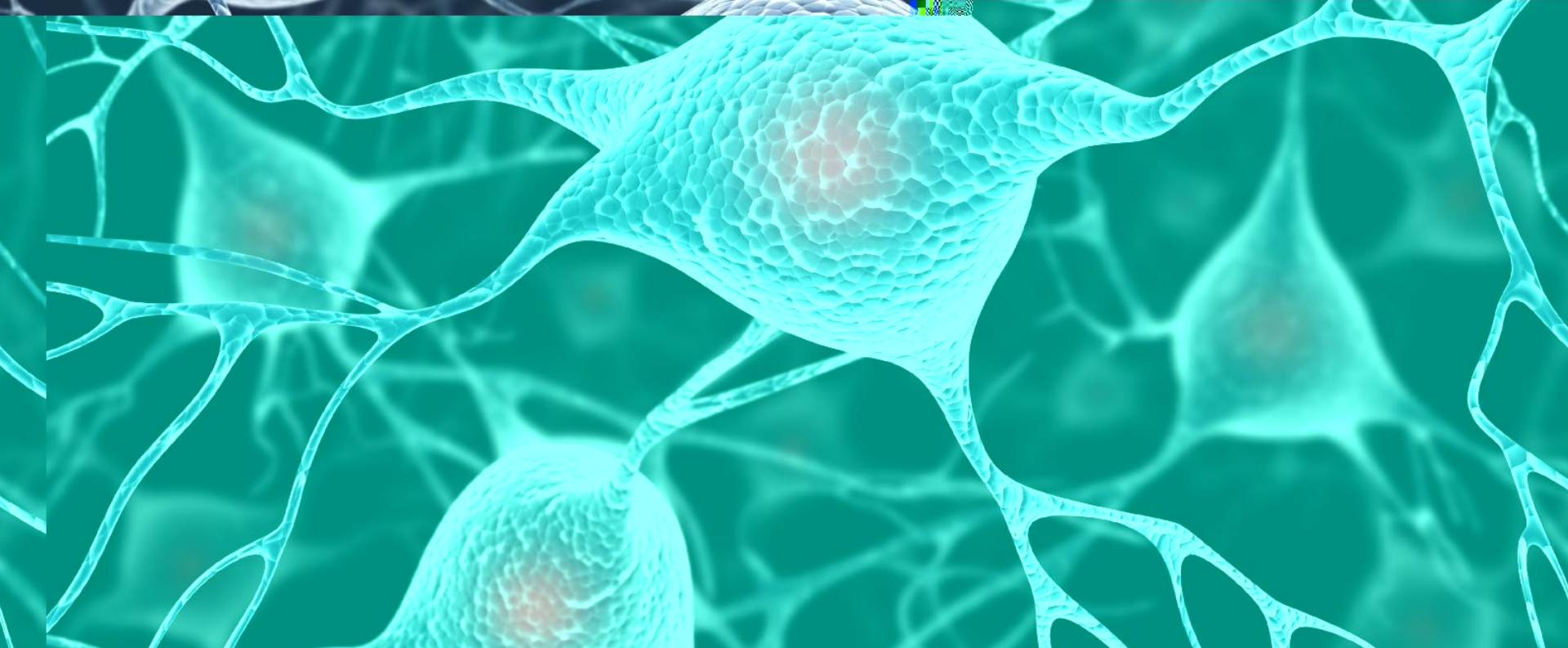
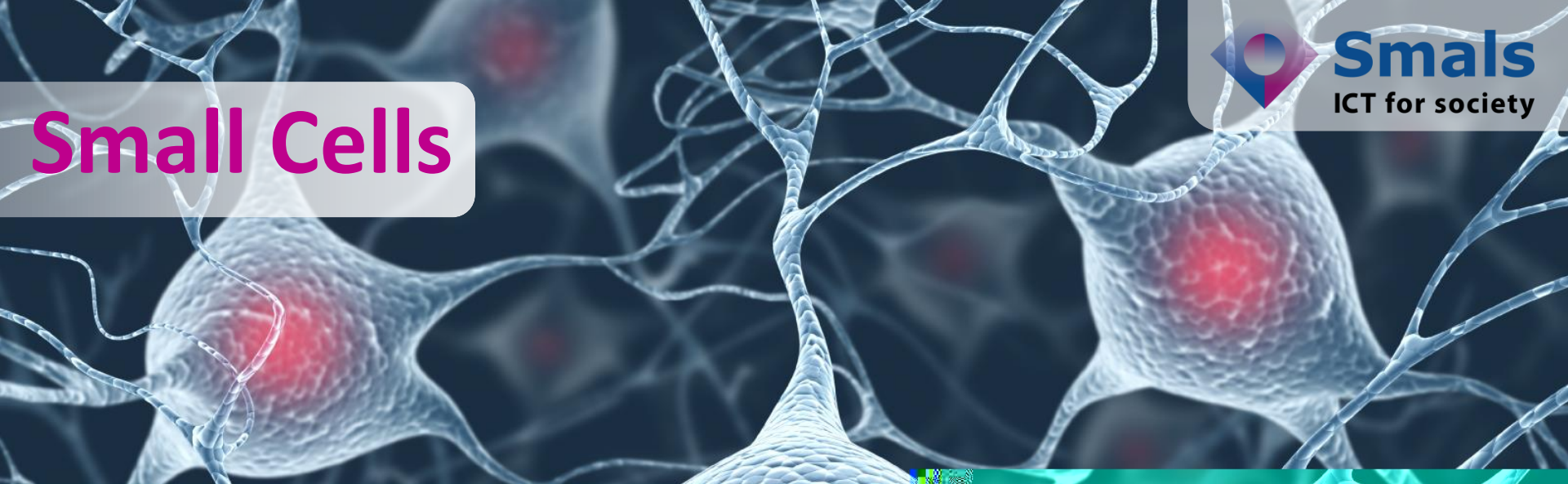
Niemand heeft een overzicht...  
Hoe lossen we dit op zonder compromitteren  
privacy burger?



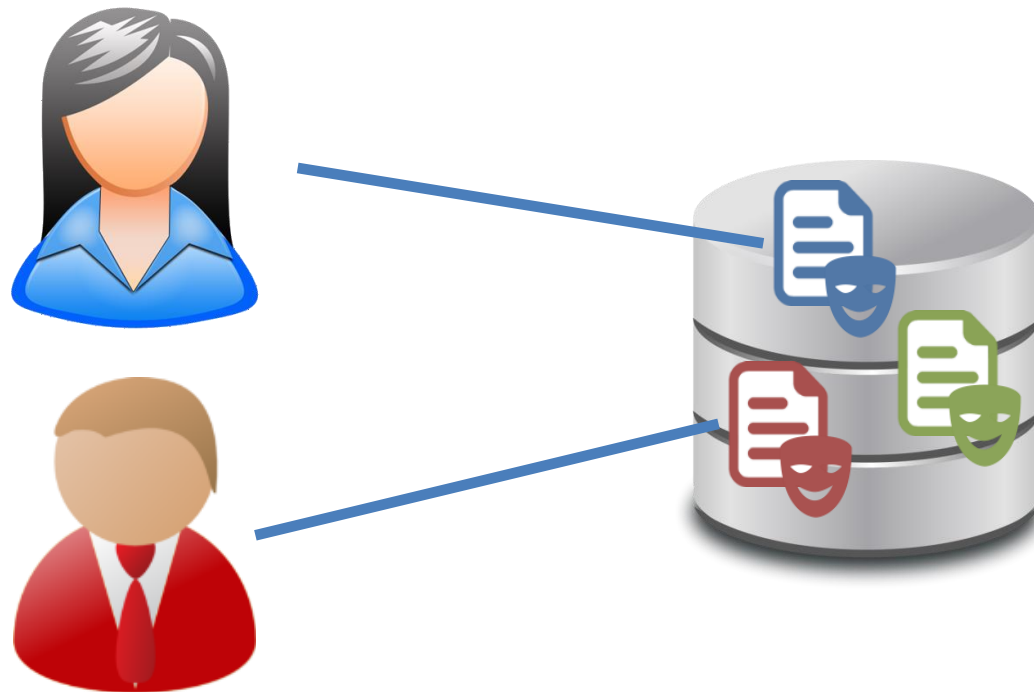
Aanvrager:	KU Leuven
Doelstelling:	Onderzoek naar jonge zelfstandigen in bijberoep in k...
Duur:	01/06/2006 tot 31/12/2016
Selectie:	YoB $\geq$ 1990 & wage > 50.000 & status=zelfst. in bijb.
Attributen:	Opleiding, gezinsinkomen, gemeente, werkgever, ...

Burger leert in welke projecten hij betrokken is.  
Overheden weten dit niet.

# Small Cells



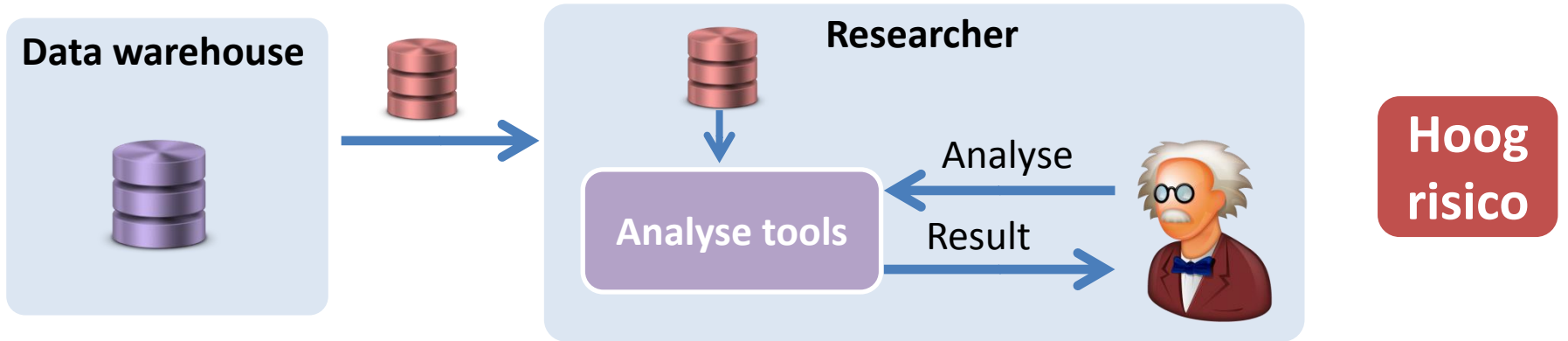
# Small Cells



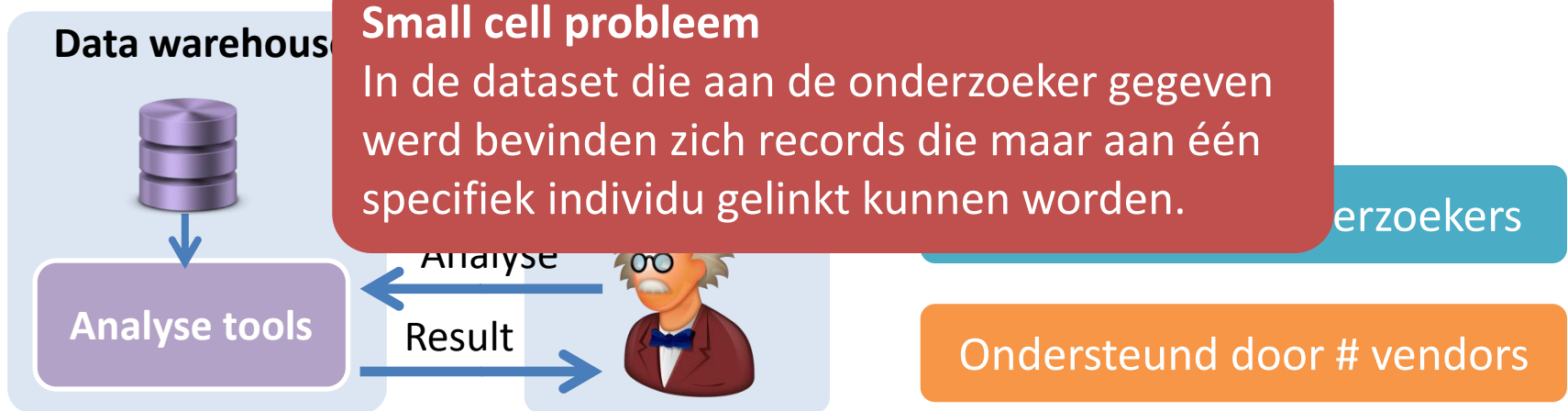
Doorsturen van “geanonimiseerde” gegevens naar onderzoeker houdt dus risico’s in

# Paradigmashift

## Haal de data naar de berekeningen

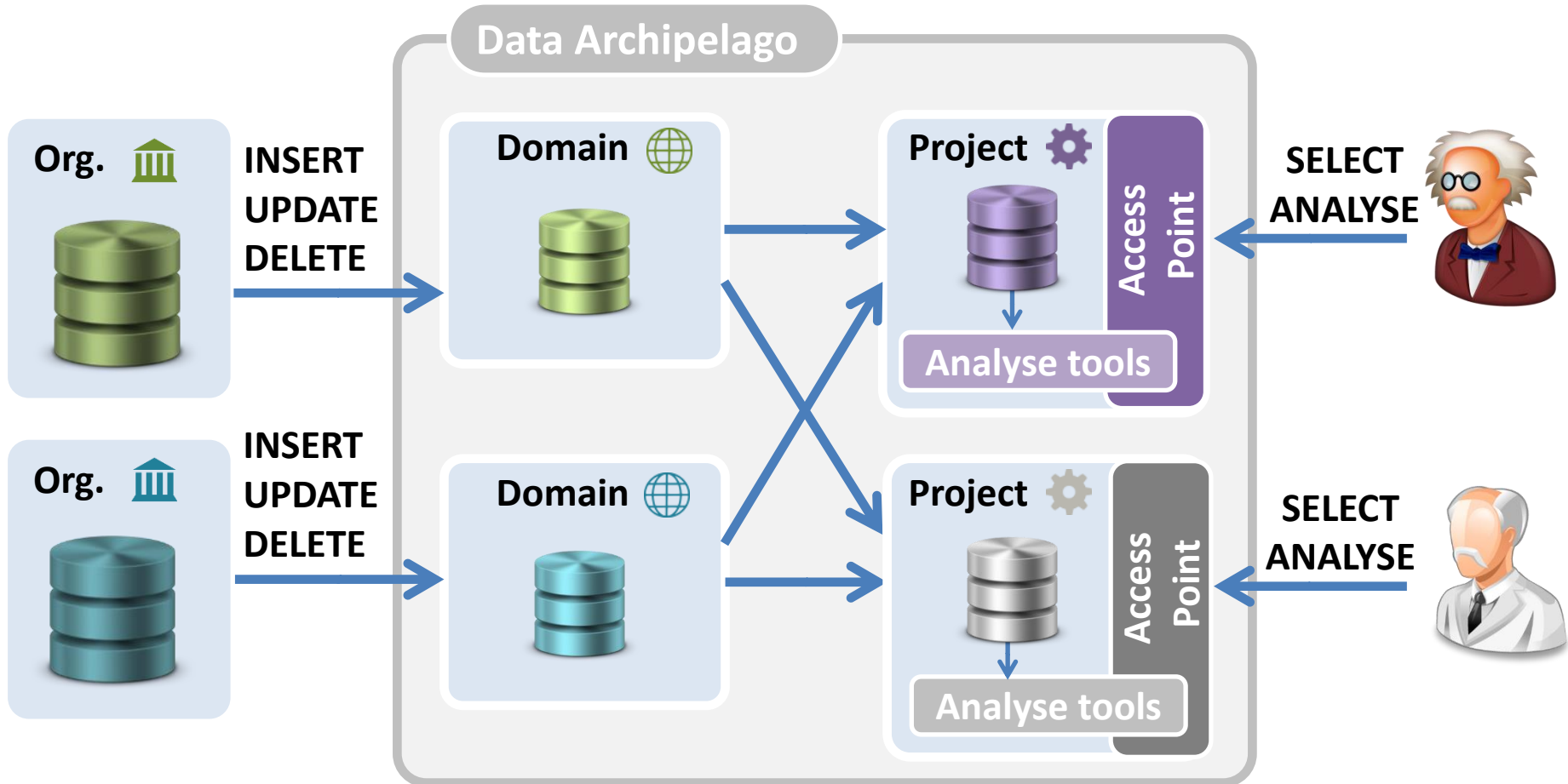


## Haal de berekeningen naar de data



**Small cell probleem**  
In de dataset die aan de onderzoeker gegeven werd bevinden zich records die maar aan één specifiek individu gelinkt kunnen worden.

# Omgaan met Small Cells



Fine-grained  
access control

Logging &  
Monitoring

Policies

Differential  
privacy

# Differential Privacy

*“Veronderstel dat je op elk moment een query kan doen op een database zodat je het gemiddelde inkomen in een wijk kent. Als je weet dat meneer Jansens naar een andere wijk verhuist en je voert de query vlak voor en vlak na zijn vertrek uit, kun je zijn loon berekenen”*

**Privacy garanties voor data  
subjecten bij recurrente  
statistische queries**

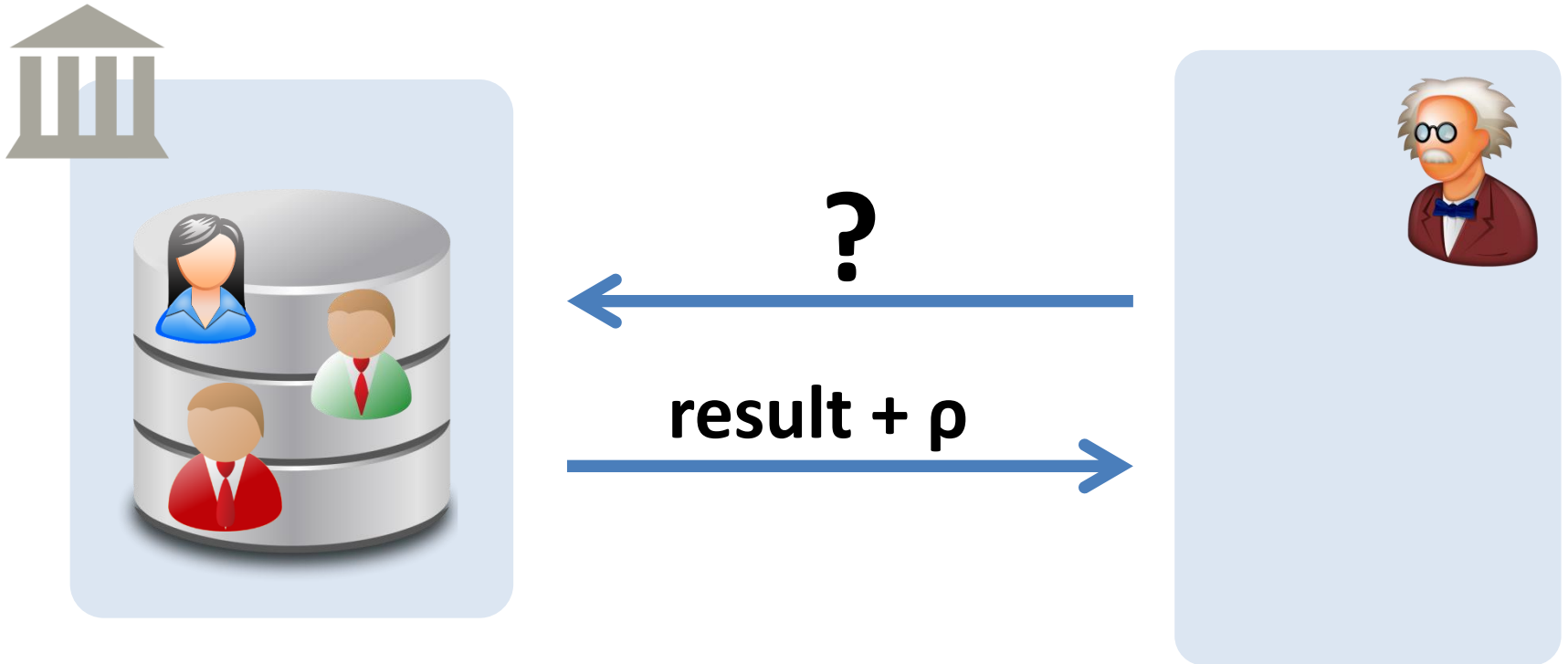
**Wiskundig  
onderbouwd**

**Vrij nieuw  
principe**

Onderzoeker kiest zelf hoeveel statistische query's hij uitvoert.

- 1x Hoge nauwkeurigheid
- 2x Lagere gemiddelde nauwkeurigheid
- 3x nog lagere gemiddelde nauwkeurigheid
- ...

# Differential Privacy



**Door nemen gemiddelde resultaat zelfde query  
kunnen we  $\rho$  grotendeels wegfilteren...**

**Een privacy budget verhindert dit.**

# Differential Privacy

## Query 1

Veel ruis, dus minder exact

$$\epsilon = \epsilon - 1$$

## Query 2

Weinig ruis, dus vrij exact

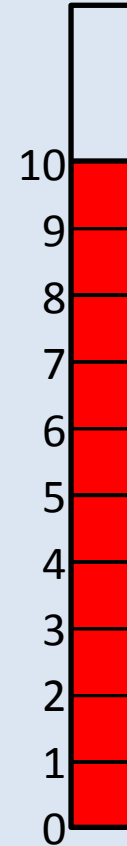
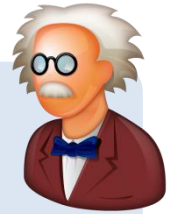
$$\epsilon = \epsilon - 6$$

## Query 3

Ruis tussenin, exactheid ook

$$\epsilon = \epsilon - 3$$

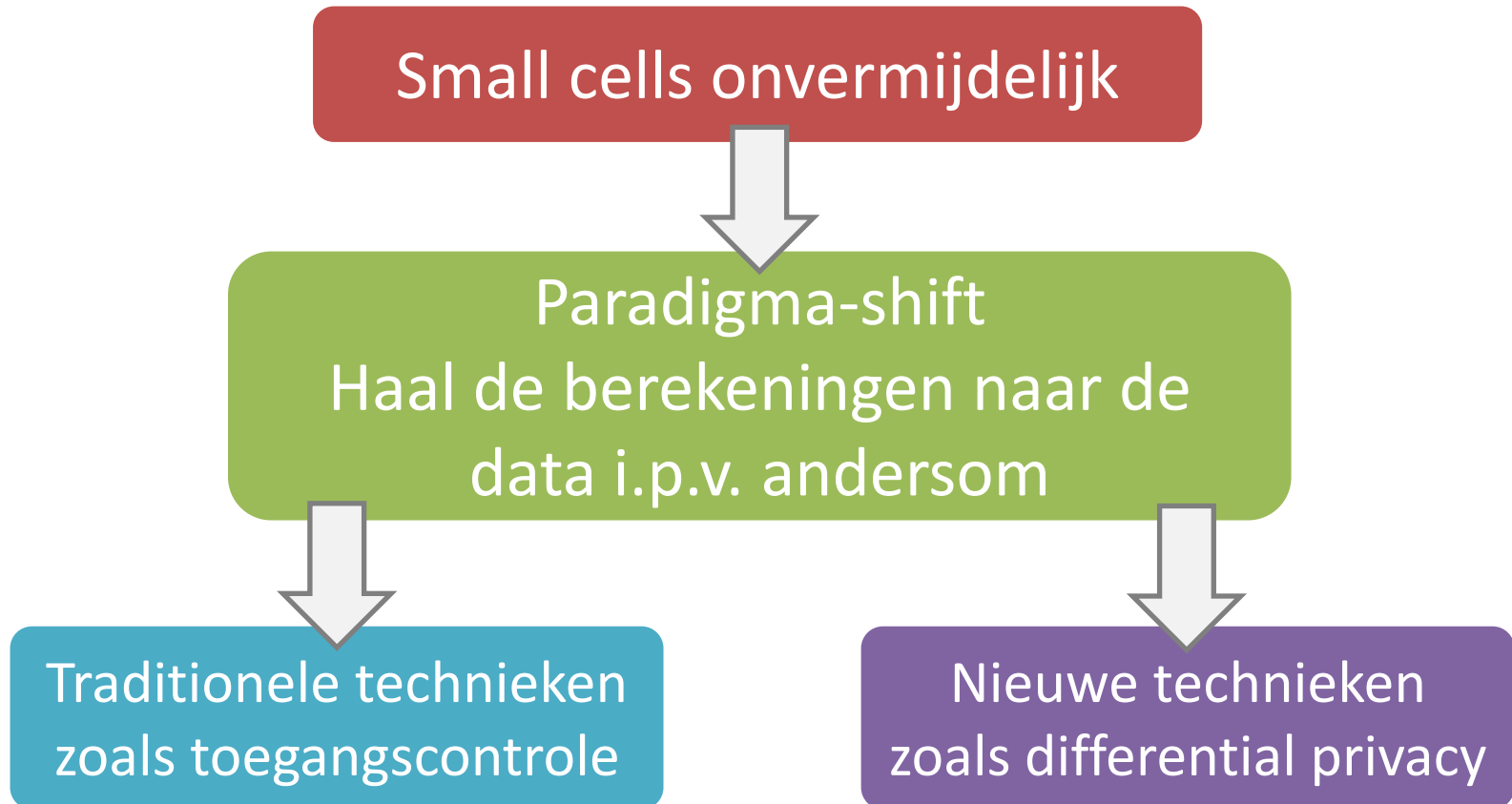
Geen verdere queries  
meer mogelijk



Privacy  
budget

$\epsilon$

# Omgaan met Small Cells



**Verder onderzoek vereist!**

# AGENDA

Introductie

Technieken Anonimisatie

Wat

Beperkingen

Data Archipel

Concept

Proof of concept

Een beetje crypto

Deanonimisatie

Sleutelbeheer

Transparantie

Small cells

Conclusies



# Conclusie & toekomst



# Kruisen van Persoonsgegevens

## Een fictief voorbeeld

*Een **onderzoeksteam** wil medische, financiële en demografische persoonsgegevens analyseren van alle **burgers** die na 1990 geboren zijn, die zelfstandige in bijberoep zijn met een loon van minstens € 50 000.*

*Deze gegevens worden echter beheerd door verschillende **overheidsbedrijven** en moeten dus gekruist worden.*

**Wetenschapper**  
Vlot kruisen data

**Burger**  
Respect  
privacywetgeving

**Overheidsorganisatie**  
Behoudt controle  
(want verantwoordelijk)

**Allen**  
minimale impact  
data breach

# Toegevoegde waarde

Combineren van

**Wetenschapper**  
Vlot kruisen data

**Burger**  
Respect  
privacywetgeving

**Overheidsorganisatie**  
Behoudt controle  
(want verantwoordelijk)

**Allen**  
minimale impact  
data breach

Binnen wettelijke kader

**Proportionaliteit**  
Naar project / deanon.

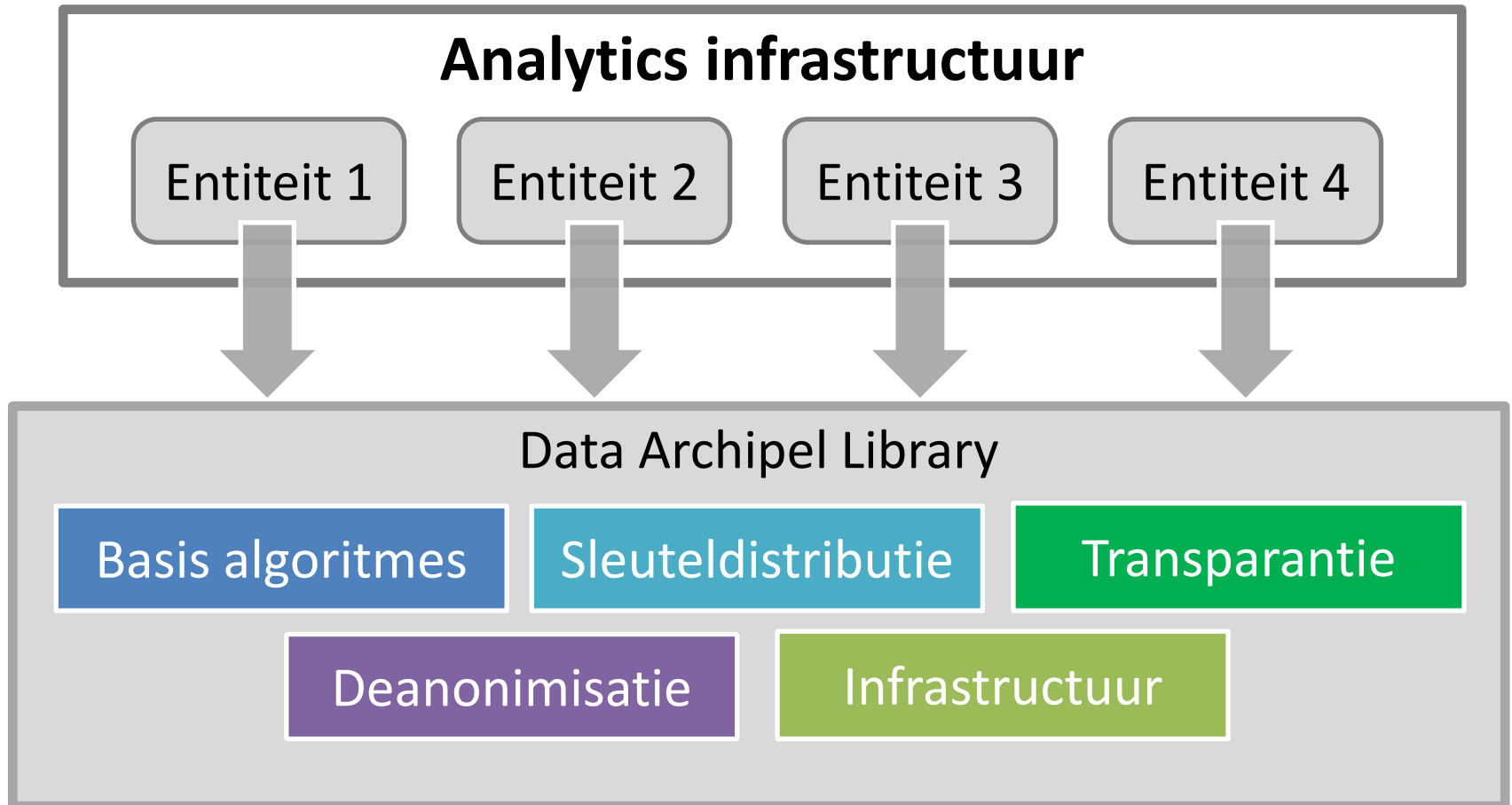
**Finaliteit**  
In "Berekeningen naar data"

**Transparantie**

**Information security practices**

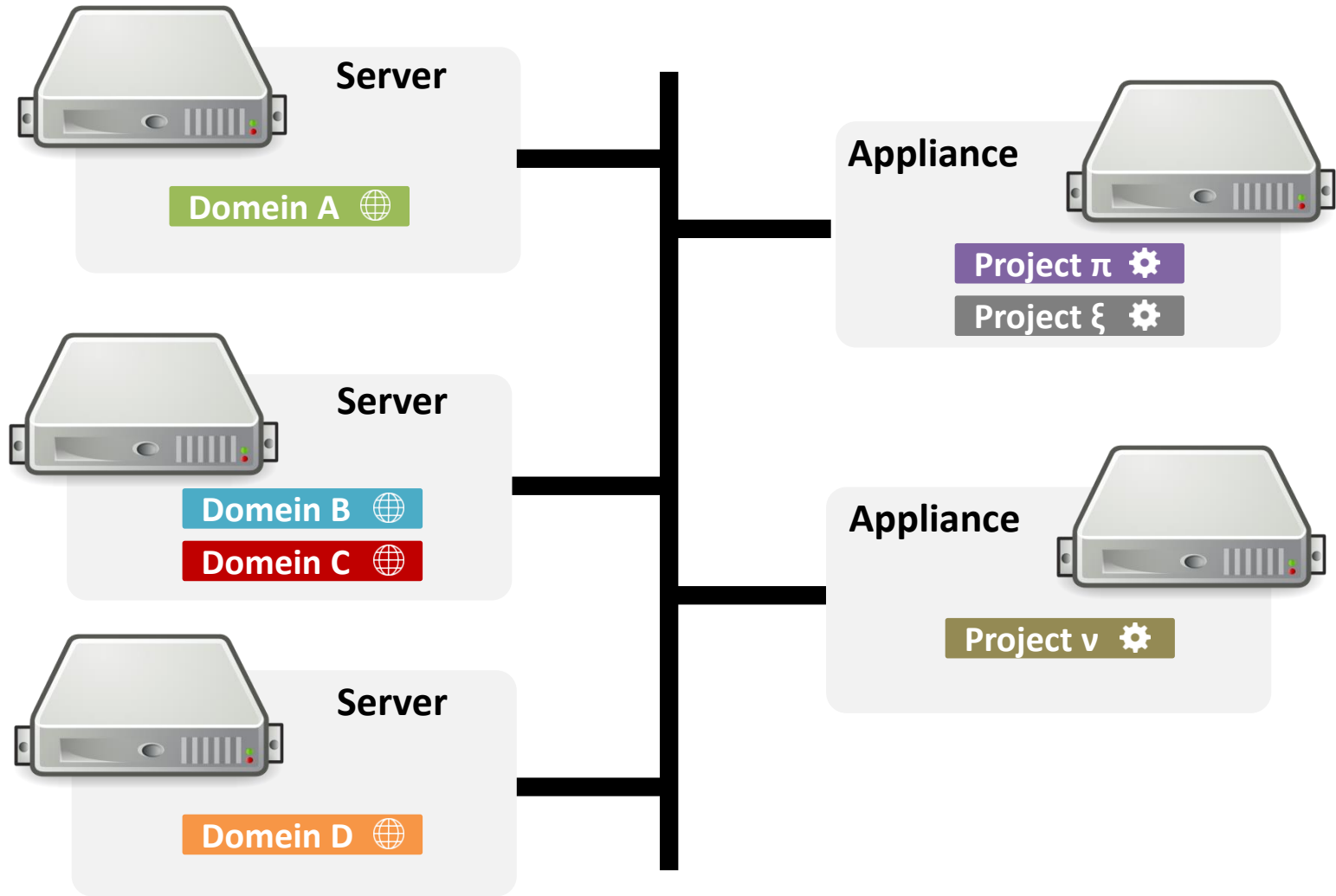
**Stap vooruit t.o.v. huidige werkwijze**

# Flexibiliteit



**GEEN monolithisch systeem**

# Gedistribueerde Infrastructuur



Horizontaal  
schaalbaar

Vereisten servers  
< appliances

[opt] zelfde DC  
→ performantie

[opt] Instelling  
beheert server

# Validatie

## Juridische validatie

- Prof. dr. Van Eecke, UA & DLA Piper

## Technische Validatie

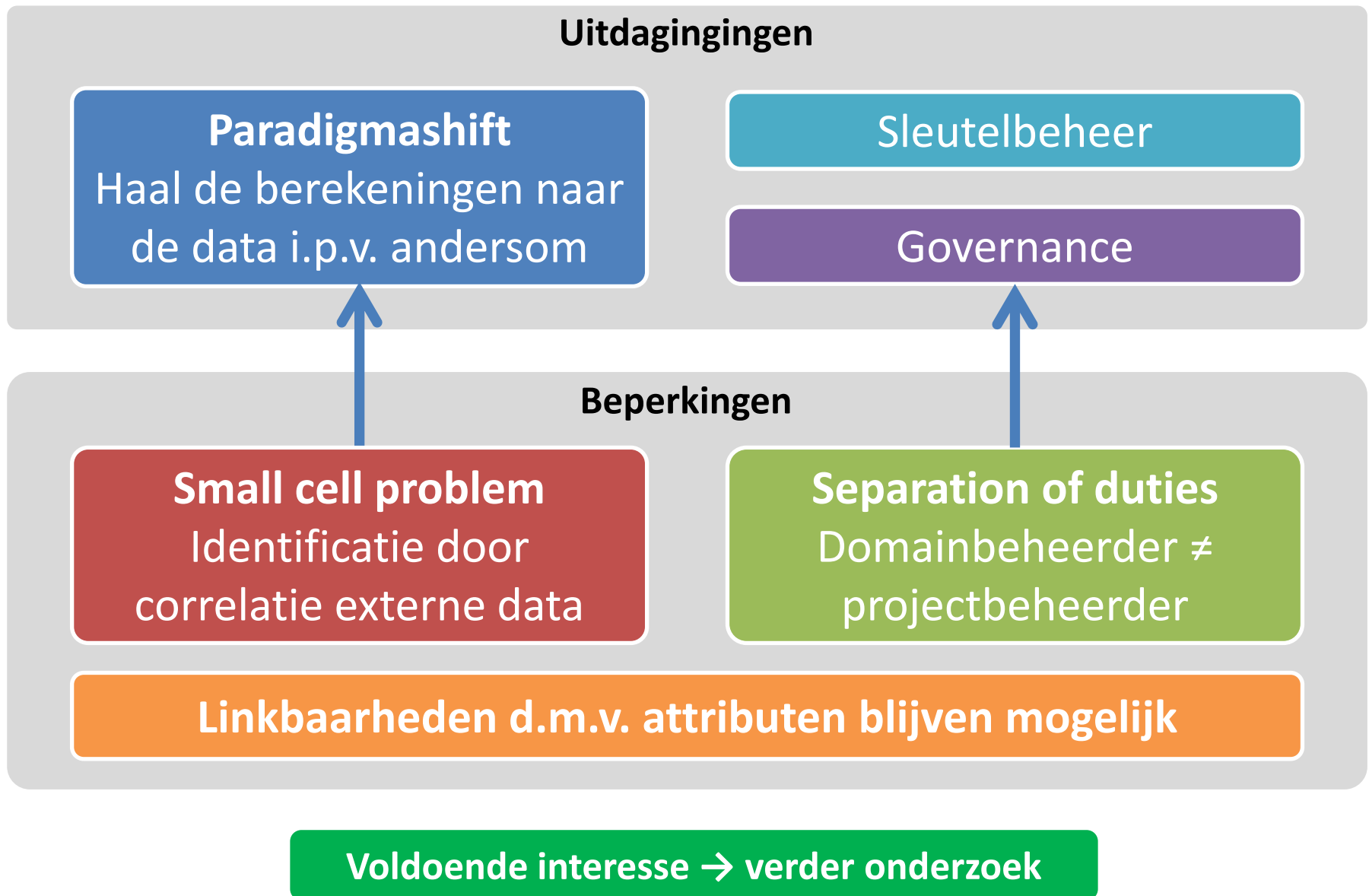
**KU LEUVEN**

- COSIC, KU Leuven (recent in WSJ)
- Departement computerwetenschappen, KU Leuven
- Internationale experts (PETS programme committee)

## Functionele validatie

- Proof-of-concept
- Concrete case: **We need you!**

# Uitdagingen & Beperkingen



# Toekomst (voorwaardelijk)

Uitbreiden &  
toepassen POC

Uptake (delen van) concept  
door vendors

Ontwikkelen visie sector  
In welke richting wat betreft  
data management bij analytics?

Realisatie  
Eventueel met privé

Verder onderzoek

Tijd



# Publicaties

## **Blogpost Anonimisatie**

*Big data & krakend ijs onder  
anonimisatie*

Mei 2015

## **Presentatie**

*Privacy vs. Analytics*

Maart 2016

## **Wetenschappelijk artikel**

*Data Archipelago -  
Reconciling privacy and  
analytics on multi-source PII*

2016

## **Toegankelijk rapport**

*Data Archipel – Analytics op  
gekruiste persoonsgegevens*

Maart 2016

[www.smalsresearch.be](http://www.smalsresearch.be)

# Crypto@gov.be

Les Crypto is een extreem krachtige, veelzijdige tool die veel verder gaat dan traditionele encryptie, veilige verbindingen en digitale handtekeningen.

**Maar onbekend = onbemind (ook in overheidscontext...)**

Data archipel

Elliptische  
krommen

Format preserving  
encryption

Bitcoin / Blockchain

Homomorphic encryption

Anonymous  
credentials

Quantum  
computing

...

**Bij voldoende interesse: Infosessie - Crypto Awareness**

## Kristof Verslype



02 787 53 76



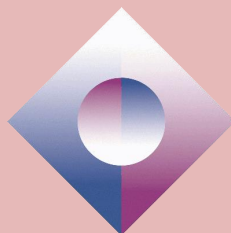
Kristof.verslype@smals.be



@KristofVerslype



## Smals



[www.smals.be](http://www.smals.be)



@Smals\_ICT



[www.smalsresearch.be](http://www.smalsresearch.be)



@SmalsResearch



# Referenties

- [AV07] Arvind Narayanan, Vitaly Shmatikov. **How To Break Anonymity of the Netflix Prize Dataset**. Cornell University Library. 22 november 2007.  
<http://arxiv.org/abs/cs/0610105>
- [B92] **Wet van 8 december 1992 tot bescherming van de persoonlijke levensfeer ten opzichte van de verwerking van persoonsgegevens**. 1999.  
<http://www.e-privacy.be/privacywet.pdf>
- [C14] Ann Cavoukian. **Big Data and Innovation, Setting the Record Straight: De-identification Does Work**. Canada. 16 juni 2014  
<http://www2.itif.org/2014-big-data-deidentification.pdf>
- [I11] Information Commissioner's Office (ICO). **Data sharing code of practice**. Mei 2011.  
[https://ico.org.uk/media/for.../1068/data\\_sharing\\_code\\_of\\_practice.pdf](https://ico.org.uk/media/for.../1068/data_sharing_code_of_practice.pdf)
- [I12] Information Commissioner's Office (ICO). **Anonymisation: managing data protection risk code of practice**. November 2012.  
<https://ico.org.uk/media/1061/anonymisation-code.pdf>
- [I14] Information Commissioner's Office (ICO). **Big data and data protection**. 28 juli 2014.  
<https://ico.org.uk/media/for-organisations/documents/1541/big-data-and-data-protection.pdf>
- [KB01] **Koninklijk besluit ter uitvoering van de wet van 8 december 1992 tot bescherming van de persoonlijke levenssfeer ten opzichte van de verwerking van persoonsgegevens**. 13 februari 2001.  
[http://www.privacycommission.be/sites/privacycommission/files/documents/CONS\\_kb\\_uitvoering\\_privacywet\\_13\\_02\\_2001.pdf](http://www.privacycommission.be/sites/privacycommission/files/documents/CONS_kb_uitvoering_privacywet_13_02_2001.pdf)
- [L01] Latanya Sweeney . k-Anonymity: A model for protecting privacy. Mei 2001.  
[https://epic.org/privacy/reidentification/Sweeney\\_Article.pdf](https://epic.org/privacy/reidentification/Sweeney_Article.pdf)

# Referenties (Ctd)

- [M13] Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen & Vincent D. Blondel. ***Unique in the Crowd: The privacy bounds of human mobility***. 25 maart 2013. Scientific Reports 3, Article number: 1376.  
[http://www.nature.com/srep/2013/130325/srep01376/fig\\_tab/srep01376\\_F1.html](http://www.nature.com/srep/2013/130325/srep01376/fig_tab/srep01376_F1.html)
- [M15a] Tania Martin, Smals research. ***Elliptic Curve Cryptography for dummies 1: introduction***. 25 februari 2015.  
<https://www.smalsresearch.be/elliptic-curve-cryptography-tutorial1/>
- [M15b] Tania Martin, Smals Research. ***Elliptic Curve Cryptography for dummies 2: en pratique pour la cryptographie***. 12 Augustus 2015  
<https://www.smalsresearch.be/elliptic-curve-cryptography-tutorial2/>
- [MSC13] Viktor Mayer-Schonberger., Kenneth Cukier. ***Big Data: A Revolution That Will Transform How We Live, Work and Think***. 10 oktober 2013.  
<http://arxiv.org/abs/1304.7605>
- [NF14] Arvind Narayanan, Edward W. Felten. ***No silver bullet: De-identification still doesn't work***. 9 juli 2014. Princeton University.  
<http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf>
- [O09] Paul Ohm. ***Broken Promises of Privacy: Responding to the surprising failure of anonymisation***. UCLA Law Review, Vol. 57, p. 1701, 2010; U of Colorado Law Legal Studies Research Paper No. 9-12. 13 augustus 2009  
[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1450006](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006)
- [PCAST14] President's Council of Advisors on Science and Technology (PCAST). ***Big Data and Privacy: A Technological Perspective***. Mei 2014.  
[http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast\\_big\\_data\\_and\\_privacy\\_-\\_may\\_2014.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf)

# Referenties (Ctd)

- [S15] Science. ***Credit card study blows holes in anonymity***. 30 januari 2015  
[http://www.sciencemagazinedigital.org/sciencemagazine/30\\_january\\_2015?folio=468#pg16](http://www.sciencemagazinedigital.org/sciencemagazine/30_january_2015?folio=468#pg16)
- [T14] Anthony Tockar. ***Differential Privacy: The Basics***. Neustar Research. 8 september 2014.  
<http://research.neustar.biz/2014/09/08/differential-privacy-the-basics/>
- [WP01] Article 29 Data Protection Working Party. ***Opinion 05/2014 on Anonymisation Techniques***. 10 april 2014  
[http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)

# Websites / Artikels

- (CNN07) CNN Money. **101 Dumbest Moments in Business**. 2007  
[http://money.cnn.com/galleries/2007/biz2/0701/gallery.101dumbest\\_2007/](http://money.cnn.com/galleries/2007/biz2/0701/gallery.101dumbest_2007/)
- (DN15) DataNews. **Anonimiteit bij big data-analyse een illusie**. 2 februari 2015.  
<http://datanews.knack.be/ict/nieuws/anonimiteit-bij-big-data-analyse-een-illusie/article-normal-529977.html>
- **(UKANON) UK Anonimisation Network**.  
<http://ukanon.net/>
- (DM15) De Morgen. **Google loopt deur plat bij Witte Huis**. 26 maart 2015.  
<http://www.demorgen.be/economie/google-loopt-deur-plat-bij-witte-huis-a2265407/>
- (NYT06) Michael Barbaro, Tom Zeller. **A Face Is Exposed for AOL Searcher No. 4417749**. 9 augustus 2006, New York Times.
- (ZD15) ZDNet. **These companies lost your data in 2015's biggest hacks, breaches**. 13 januari 2016.  
<http://www.zdnet.com/pictures/worst-largest-security-data-breaches-2015/13/>
- (OPM) Wikipedia. **Office of Personnel Management data breach**.  
[https://en.wikipedia.org/wiki/Office\\_of\\_Personnel\\_Management\\_data\\_breach/](https://en.wikipedia.org/wiki/Office_of_Personnel_Management_data_breach/)
- (PETS) Privacy Enhancing Technologies Symposium.  
<https://www.petsymposium.org/>
- (WSJ) Wall Street Journal. **In Belgium, an Encryption Powerhouse Rises**. 10 December 2015  
<http://www.wsj.com/articles/in-belgium-an-encryption-powerhouse-rises-1449791014>
- (JZ) Jianying Zhou. **Top Crypto and Security Conferences Ranking (2015)**.  
<http://icsd.i2r.a-star.edu.sg/staff/jianying/conference-ranking.html>