

Research dataplatform

ELSI workshop gegevensbescherming - 19 feb 2021

Wietze Pasma

PhD "hergebruik anesthesiedata voor onderzoek" (07-2020)

Product Owner data voor onderzoek (DIT), UMC Utrecht

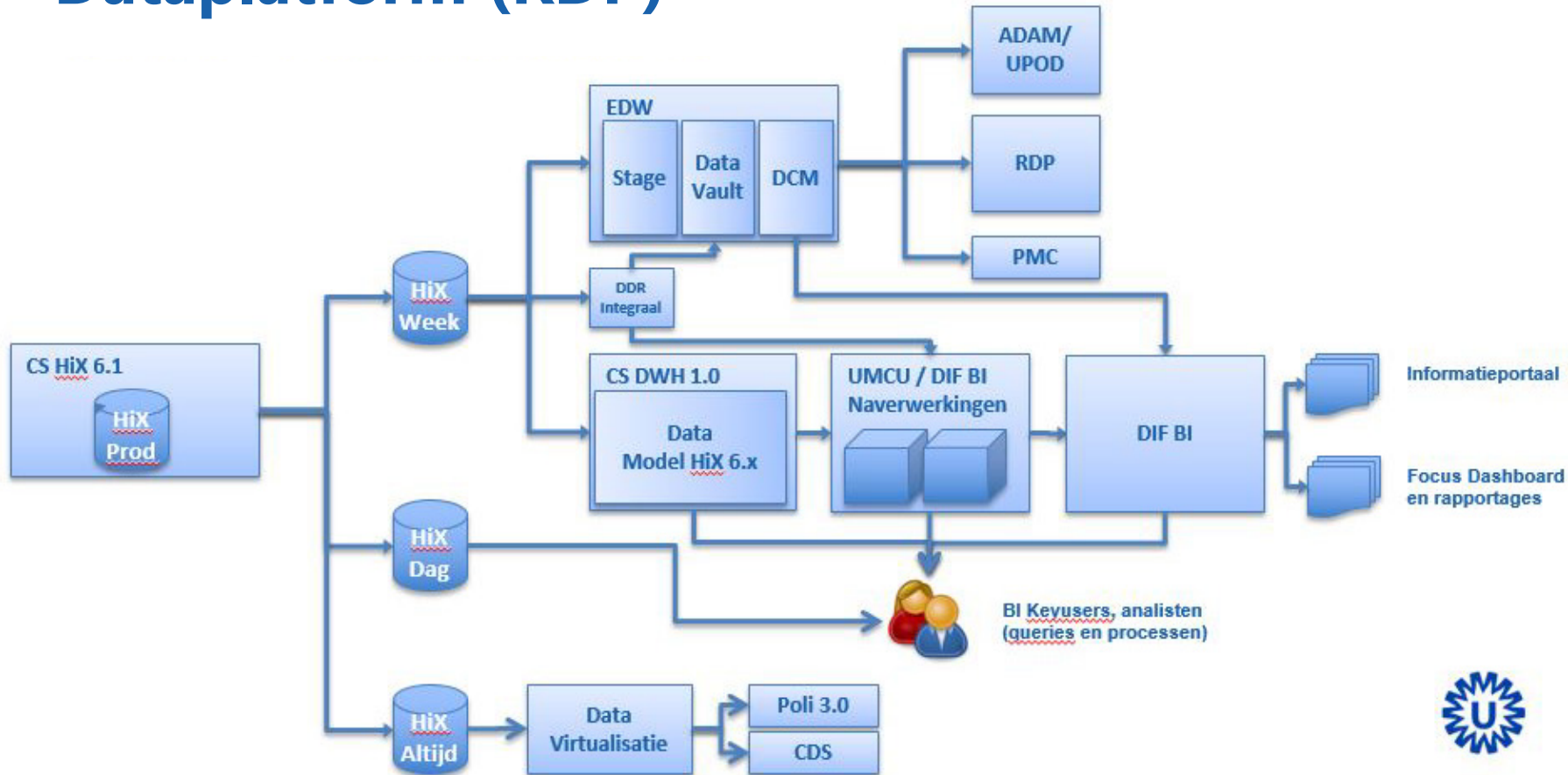
Inhoud

- UMC Utrecht dataplatform (RDP)
- Pseudonimisatie
- Nieuwe dataplatform

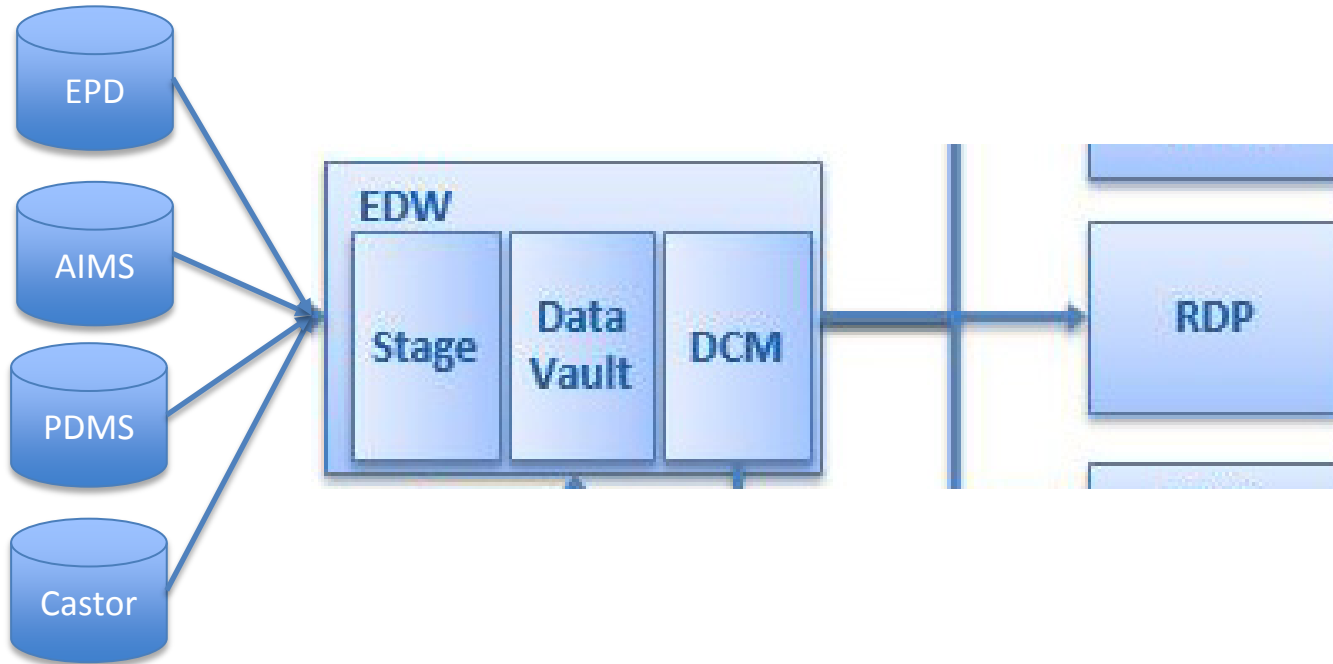
Dataplatform UMCU – waarom?

- Gestart in 2012 (project ResearchIT)
- Hergebruik Data
- Niet meer overtypen
- Privacy
- Kwaliteit
- Efficiëntie

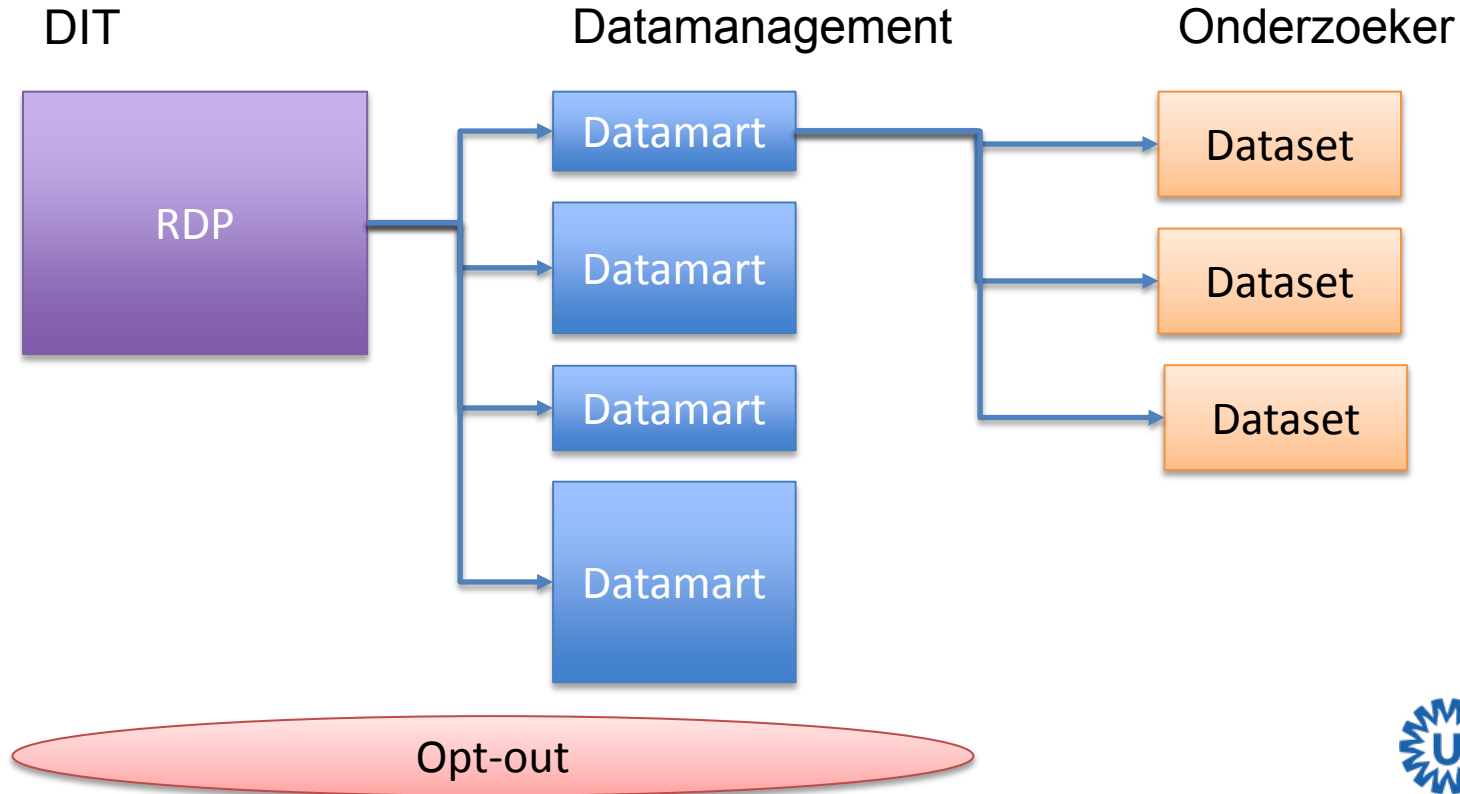
Dataplatform (RDP)



Dataplatform (RDP)



RDP data uitgifte



Eenvoudige pseudonimisatie

Situatie:

- Per patientID 1 uniek pseudoID, over alle datamarts heen

PATIENT_ID	Uitgifte_Kenmerk	PseudoID_Uitgifte
1	BIOBANK_LMS	3454
1	DIGD	3454
1	DLA	3454
1	ECZEEM	3454
1	GENEESMIDDELENCIE	3454
1	INFECTIE	3454

Gevolg:

- Datamart gebruikers kunnen middels pseudoID gegevens in verschillende datamarts aan elkaar **koppelen**
- Losse uitgiftes op basis van pseudoID zijn aan andere datamarts en/of andere uitgiftes te **koppelen**

Eenvoudige pseudonimisatie

Nadeel koppelen datamarts:

- Is niet AVG proof
- PseudoID is naast herleidbaar ook identificerend geworden, en daarmee vrijwel gelijk aan een patiëntennummer
- Geen controle of de data voor de betreffende onderzoeksvraag volledig is
- Medewerkers/onderzoekers krijgen toegang tot data van buiten hun eigen afdeling/studie, waarvoor geen toestemming is gegeven
- Maakt registratie naast de bron mogelijk

Dubbele pseudonimisatie

Nieuwe situatie:

- Per datamart uniek pseudoID

PATIENT_ID	Uitgifte_Kenmerk	PseudoID	PseudoID_DM
1	BIOBANK_LMS	3454	5649
1	DIGD	3454	543
1	DLA	3454	67547
1	ECZEEM	3454	76
1	GENEESMIDDELENCIE	3454	989
1	INFECTIE	3454	6578

Gevolg:

- Datamarts en datasets niet meer onderling te koppelen zonder tussenkomst datamanagement
- Data uitgegeven vanuit dezelfde datamart wel onderling te koppelen

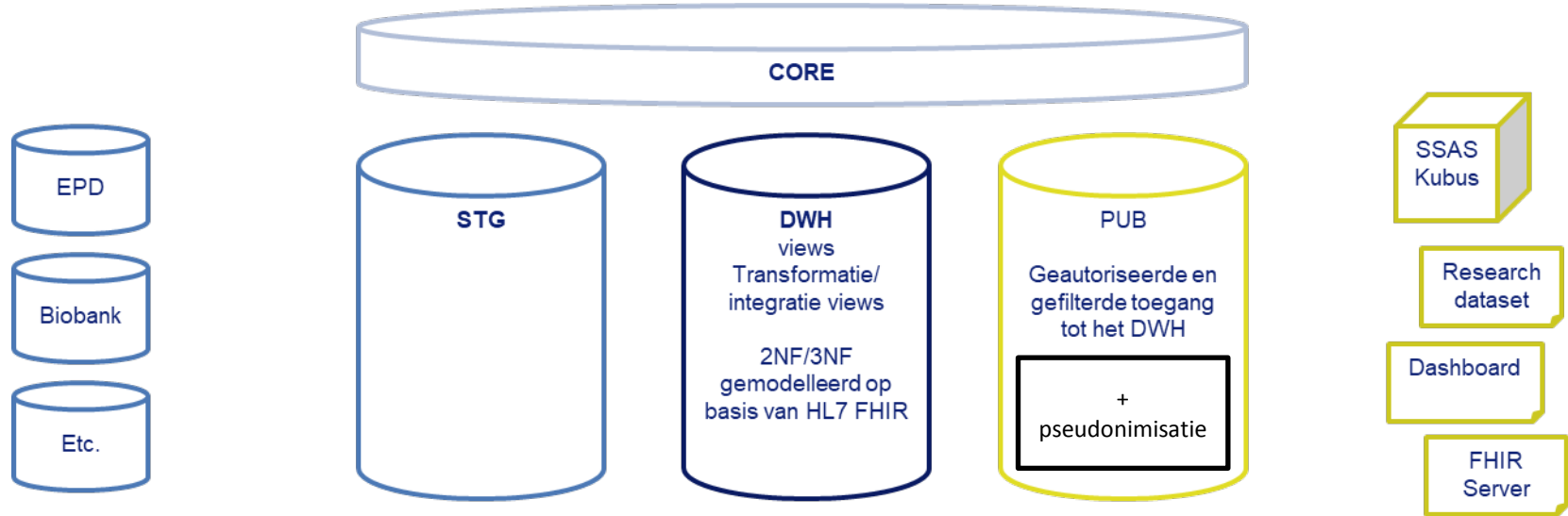
Data uitgifte

- HiX nummers uitgeven
- PseudoID uitgeven
- Eigen pseudonimisatiestap toepassen
- Anonieme data uitgeven (verschillende niveaus)

Nieuwe dataplatform

- Real time data
- FHIR datamodel, i.p.v. DCM
- Data uitwisseling mogelijk
- Meerdere centra met dezelfde architectuur (EMC, LUMC)
- Geen uitgebreide weekendverwerking
- Naast onderzoek ook toepassing in zorg

Nieuwe dataplatform



Nieuwe dataplatform

- Per uitgifte uniek pseudoID

PATIENT_ID	Uitgifte_Kenmerk	PseudoID_DM
1	BIOBANK_LMS	100126478
1	DIGD	100134891
1	DLA	100122456
1	ECZEEM	100148921
1	GENEESMIDDELENCIE	100942477
1	INFECTIE	100026312

Vragen?

