



Logius  
*Ministerie van Binnenlandse Zaken en  
Koninkrijksrelaties*

## **Versiebeheer in de Stelselcatalogus**

*Gerard Kuys, vrijdag 26 februari 2016*



## **Backlog Stelselcatalogus (versie 0.7, 02-12-2015):**

- “Uitbreiding van de content (dus nieuwe BR/Laan) moet makkelijk aangesloten kunnen worden.”
- “Ik wil als afnemer 'in de geschiedenis' kunnen kijken. Na een update van de content moet ook de oude content nog beschikbaar blijven.”



## Opgehaald bij stakeholders:

- Gebruiksvriendelijker
- Betere informatie voor de gebruiker
- Betere zoekfunctie
- Intelligentener
- Notificatie bij wijzigingen
- Goede doorverwijzingen naar de relevante organisaties
- Koppeling met LiDO van KOOP



Opgehaald bij stakeholders:

- Andere gegevensbronnen: stelsel van overheidsgegevens
- Gebeurtenissen
- Processen
- Wie deelt welke gegevens met wie: gegevensstromen
- **Versiebeheer**

Let op



Het volgende over versiebeheer in de Stelselcatalogus is (nog) geen besluit, maar een denkrichting!!



# Hoe gaan we dat doen?



Je moet kunnen vastleggen:

- Welke data toegevoegd, gewijzigd of verwijderd zijn
- Wat de wijziging inhoudt (van oud naar nieuw)
- Wie precies deze wijzigingen heeft doorgevoerd en namens wie
- Op welk moment de wijziging werd doorgevoerd
- Of dat moment ook het moment is waarop de nieuwe inhoud officieel geldig is

# Hoe gaan we dat doen?



- Je moet kunnen nagaan:
  - Of de wijziging effect heeft (gehad) op andere data, welk effect dat was, en wanneer precies
- Je moet kunnen reconstrueren hoe de situatie op een gegeven tijdstip was
  - Implementeren van het Memento-pattern (de verzameling van voorafgaande toestanden van het systeem)

# De eindige-toestands automaat



WIKIPEDIA  
De vrije encyclopedie

- Hoofdpagina
- Vind een artikel
- Vandaag
- Etalage
- Categorieën
- Recente wijzigingen
- Nieuwe artikelen
- Willekeurige pagina

- Informatie
- Gebruikersportaal
  - Snelcursus
  - Hulp en contact
  - Donaties

Niet aangemeld Overleg Bijdragen Registreren Aanmelden

Artikel [Overleg](#)

Lezen

[Bewerken](#)

[Geschiedenis](#)

Zoeken



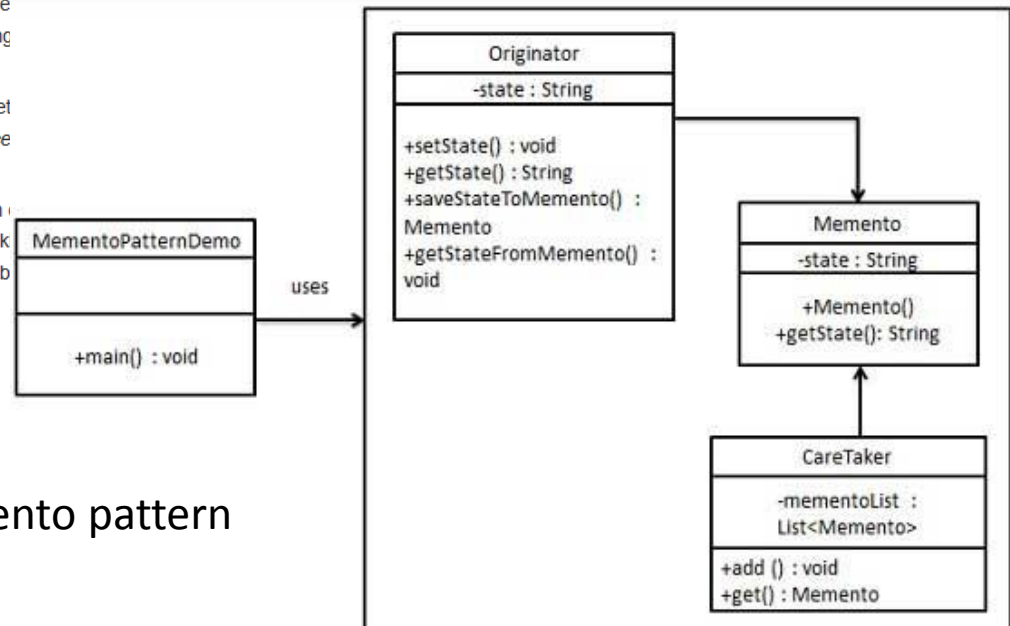
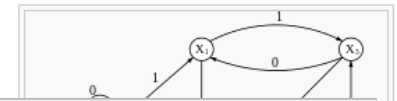
## Eindigetoestandsautomaat

Een **eindigetoestandsautomaat** (in het Engels: finite-state automaton veelal afgekort tot FA) is een abstract, wiskundig model voor het gedrag van een systeem waarbij het model bestaat uit een eindig aantal toestanden, overgangen tussen die toestanden en acties.

In de tekening hiernaast is een FA te zien met vijf toestanden  $X_0, \dots, X_4$ , aangegeven met cirkels, tien overgangen tussen de vijf toestanden elk aangegeven door een doorgetrokken lijn met de doorgetrokken gerichte lijnen tussen de toestandsovergang toestand, dit zijn *lussen* ofwel *self-loops*.

De begintoestand wordt vaak aangegeven met een uit het niet dubbele cirkel. In de tekening is  $X_0$  de begintoestand. De acceptatie toestand (zie verderop) in de figuur is  $X_4$ .

Eindigetoestandsautomaten worden onder meer toegepast in de computerwetenschap, worden vaak gerepresenteerd als [graf](#)en en beschrijven een klein deel van de natuurkunde, biologie en taalkunde, modellen ontwikkeld die gebaseerd zijn op de theorie van de eindigetoestandsautomaten.



heeft als implementatie: het Memento pattern



# Hoe doe je dat met Linked Data?

Be the first to clip this slide

## Our Goal

- Integrate provenance metadata into the Web of data to enable information quality assessment
- Vocabulary to describe provenance of Linked Data on the Web
- Easy to use
  - By people who provide Linked Data
  - By developers of Linked Data publishing tools

May 14, 2010

The Provenance Vocabulary

2

# Het PROV raamwerk



- PROV staat voor Herkomst (*Provenance*)
- Bedoeld om vast te leggen waar wijzigingen vandaan komen
- En welke wijzigingen dat zijn geweest
- Toe te passen voor alle gevallen waarin het gebruik van `dct:source` niet volstaat

| Term Name: source    |   |
|----------------------|---|
| <b>URI:</b>          | <a href="http://purl.org/dc/terms/source">http://purl.org/dc/terms/source</a>   |
| <b>Label:</b>        | Source  |
| <b>Definition:</b>   | A related resource from which the described resource is derived.  |
| <b>Comment:</b>      | The described resource may be derived from the related resource in whole or in part. Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.   |
| <b>Type of Term:</b> | <a href="#">Property</a>  |
| <b>Refines:</b>      | <a href="http://purl.org/dc/elements/1.1/source">http://purl.org/dc/elements/1.1/source</a>   |
| <b>Refines:</b>      | <a href="http://purl.org/dc/terms/relation">http://purl.org/dc/terms/relation</a>   |
| <b>Version:</b>      | <a href="http://dublincore.org/usage/terms/history/#sourceT-001">http://dublincore.org/usage/terms/history/#sourceT-001</a>   |
| <b>Note:</b>         | This term is intended to be used with non-literal values as defined in the DCMI Abstract Model ( <a href="http://dublincore.org/documents/abstract-model/">http://dublincore.org/documents/abstract-model/</a> ). As of December 2007, the DCMI Usage Board is seeking a way to express this intention with a formal range declaration. |



- PROV staat voor Herkomst (*Provenance*)
- Bedoeld om vast te leggen waar wijzigingen vandaan komen
- En welke wijzigingen dat zijn geweest
- Toe te passen voor alle gevallen waarin het gebruik van `dct:source` niet volstaat
  
- Het PROV raamwerk bestaat onder meer uit:
  - PROV –DM (het datamodel)
  - PROV-O (het datamodel omgezet naar een ontologie in OWL)
  - PROV-CONSTRAINTS (een verzameling restricties die het model logisch consistent ('geldig') maken)
  - PROV-AQ (gestandaardiseerde manieren om een PROV-dataset te benaderen en te bevragen)

# Het PROV raamwerk



(<https://www.w3.org/TR/2013/NOTE-prov-overview-20130430/>)

## Bijdragen van onder anderen:

Yolanda Gil



Kai Eckert



Paul Groth



Paul Groth

### Provenance and Trust

Provenance refers to the origins of objects, otherwise influencing that object. By knowing where an object was generated and can validate that it was generated according to the PROV model.

...cripts to Europeana (DM2E) Data workflows and data enrichment

...s: Big data extraction from the Web.

I research the integration and use of diverse contextualized information with a particular emphasis on data provenance.



- Beschouw een entiteit vanuit haar life cycle
  - Dan wordt ze gegenereerd, gewijzigd en evt. verwijderd
- Elke stap in deze cyclus is een activiteit die wordt uitgevoerd door iets of iemand op een bepaald moment
  - Die iets of iemand kan dit namens iets of iemand anders doen



- Beschouw een entiteit vanuit zijn life cycle
  - Dan wordt ze gegenereerd, gewijzigd en evt. verwijderd
- Elke stap in deze cyclus is een activiteit die wordt uitgevoerd door iets of iemand op een bepaald moment
  - Die iets of iemand kan dit namens iets of iemand anders doen
- Verschillende 'agents' kunnen verschillende wijzigingen doorvoeren
- Wanneer die elkaar tegenspreken ontstaan *entailment regimes*
- Maar dit gaat voor de Stelselcatalogus vooralsnog te ver

# Waarom is dit geschikt voor versiebeheer?

- We kunnen volgen hoe, door wie en wanneer wijzigingen worden doorgevoerd
- We kunnen vanaf dat moment dus ook voorbijgestane toestanden van het systeem terughalen

## Versiebeheer



| Versiebeheer        |              |                      |                       |
|---------------------|--------------|----------------------|-----------------------|
| <i>Versienummer</i> | <i>Datum</i> | <i>Verstuurd aan</i> | <i>Opmerkingen</i>    |
| 0.1                 | 01-01-2010   |                      | Eerste opzet          |
| 0.2                 | 10-02-2010   |                      | Na opmerkingen<br>... |
| ....                |              |                      |                       |
| 1.0                 | 11-11-2010   |                      | Definitief            |

# Waarom is dit geschikt voor versiebeheer?

- We kunnen volgen hoe, door wie en wanneer wijzigingen worden doorgevoerd
- We kunnen vanaf dat moment dus ook voorbije toestanden van het systeem terughalen

## Maar:

- Op dit moment zijn in de Stelselcatalogus de metadata per basisregistratie gescheiden, zodat wijzigingen in de ene basisregistratie geen effect hebben op de metadata van de andere
- Hoe zal het gaan wanneer zich gemeenschappelijke noties van een begrip uitkristalliseren, over basisregistraties heen?
- En zullen externe datasets, waarnaar de Stelselcatalogus doorlinkt, altijd wel hetzelfde begrip onder dezelfde URI houden?



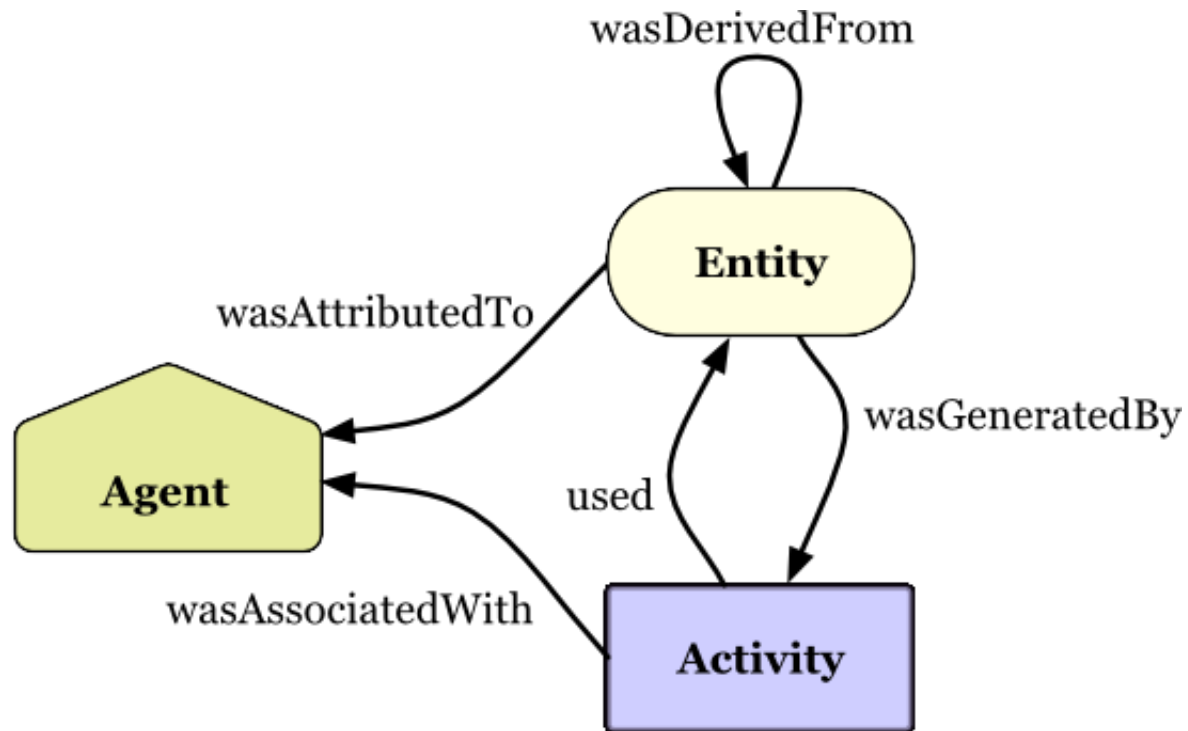
# Drie sterke pijlers onder de nieuwe Stelselcatalogus



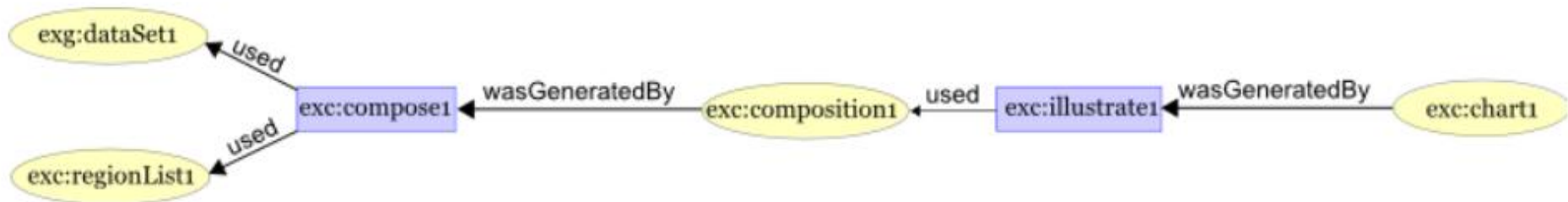
- Nog steeds geldt: geen centraal, maar een federatief model
- Versiebeheer legt de verantwoordelijkheid voor (de impact van) wijzigingen veel meer bij de bronhouder
- Versiebeheer (ook van metadata!) kan worden doorgevoerd zowel op het gebied van de materiële\* als ook van dat van de formele\* historie
  - zodat beide aspecten principieel te onderscheiden zijn

\* Materiële historie : wanneer is een feit vastgesteld

\* Formele historie : wanneer is een besluit bekend gemaakt / van kracht geworden

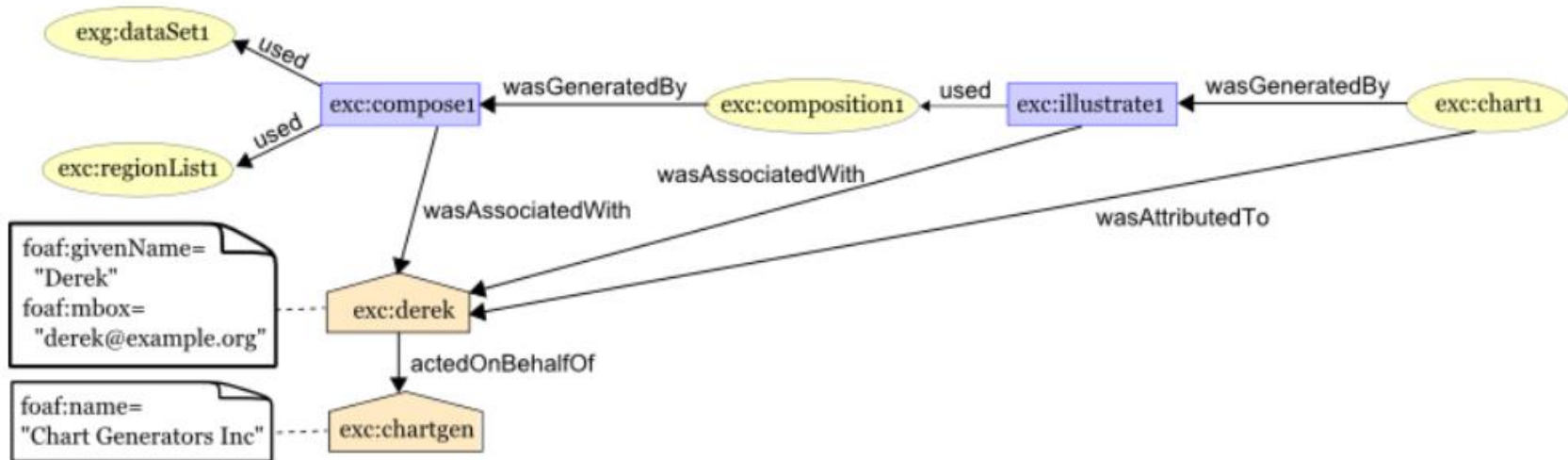


# Zo ziet een PROV-herkomstketen er uit



Van rechts naar links: Hoe een grafiek wordt gegenereerd uit twee datasets.  
Wie dat doet? Niet vastgelegd!

# Zo ziet een PROV-herkomstketen er uit

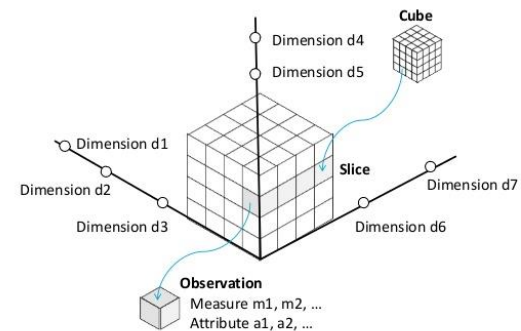


Van rechts naar links: Hoe een grafiek wordt gegenereerd uit twee datasets.  
Wie dat doet? Derek van de firma Chart Generators!

# PROV toegepast op overheidsdata

- Wat voor PROV-entiteiten
  - Versies van beleidsnota's
  - Versies van vergunningen
  - Vaststellingen van inkomen
  - Herzieningen van metadata

## RDF Data Cube 101 - Slices and observations



8 | Using the Data Cube vocabulary for Publishing Environmental Linked Data on [lab.environment.data.gov.au](http://lab.environment.data.gov.au) | Laurent LeFort

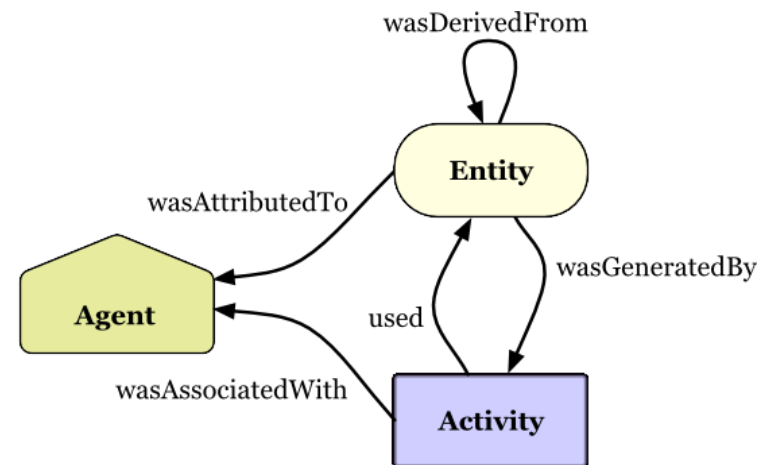


# PROV toegepast op overheidsdata

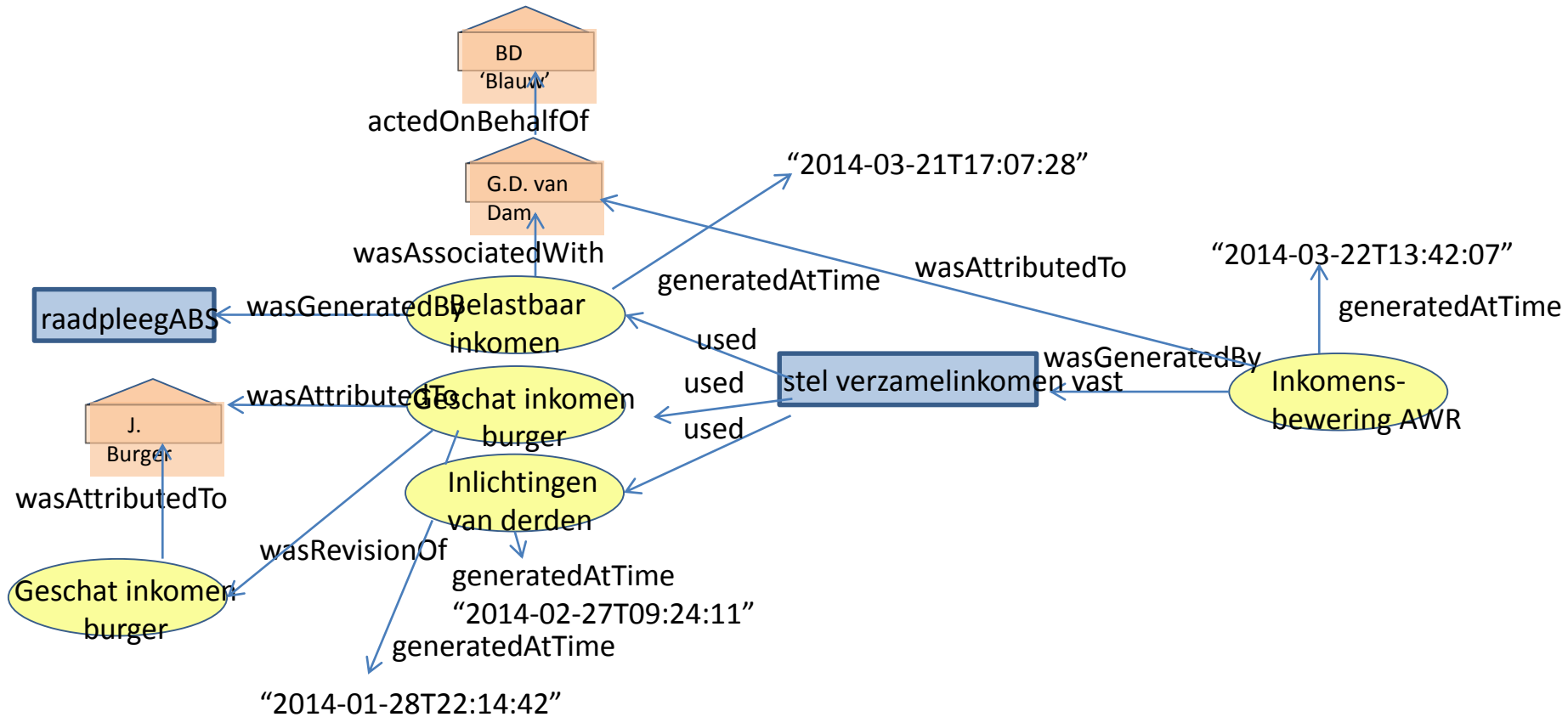
- Wat voor PROV-entiteiten
  - Versies van beleidsnota's
  - Versies van vergunningen
  - Vaststellingen van inkomen
  - Herzieningen van metadata en informatiemodellen

- Wat voor PROV-activiteiten
  - Besluiten
  - Waarnemingen

- Wat voor PROV agents
  - Wettelijk verantwoordelijken
  - Overheidsdienaren
  - Terugmelders
  - Aanvragers
  - Indieners van beroep / bezwaar



# PROV toegepast op overheidsdata: BRI



# PROV en abstractieniveaus



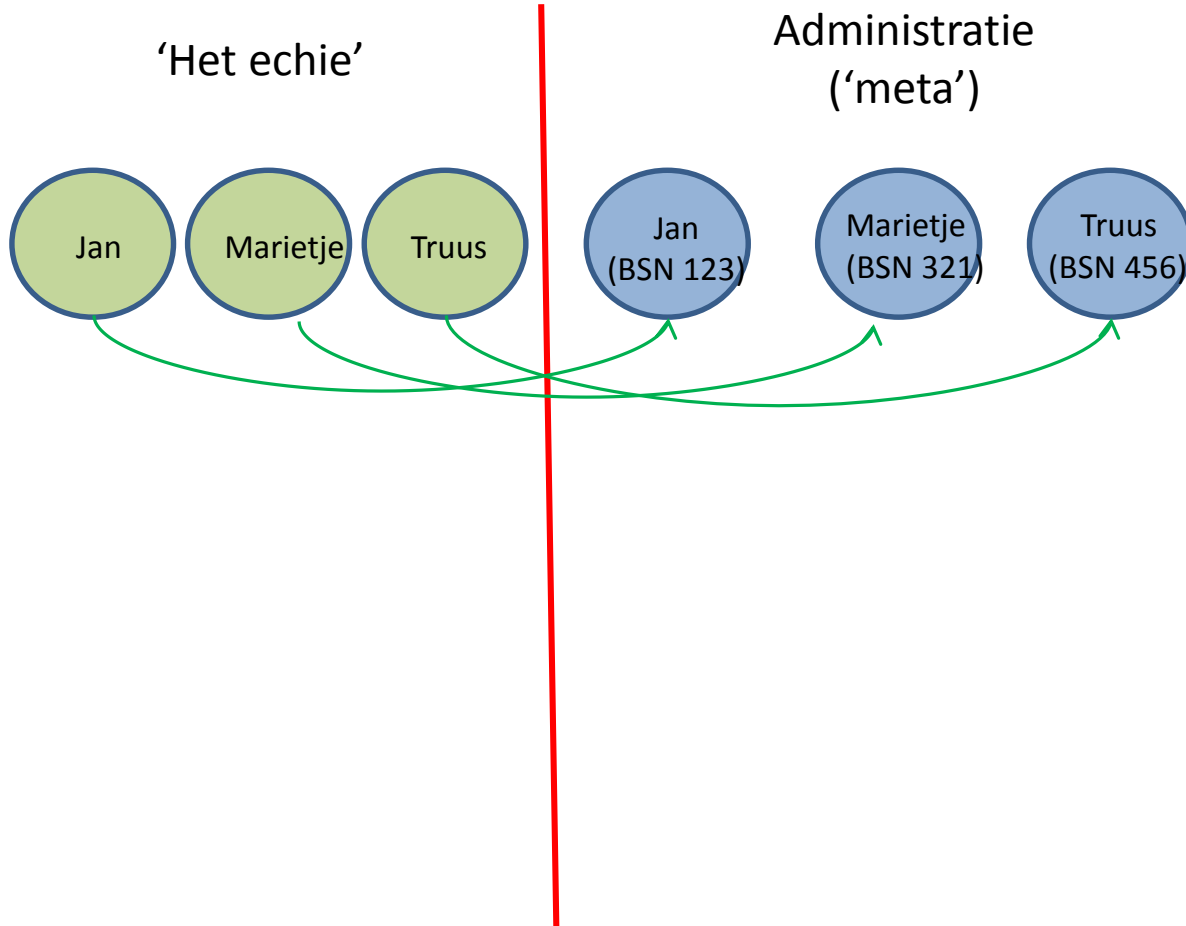
'Het echie'

Administratie  
(‘meta’)

Procesverloop  
administratie  
(‘metameta’)

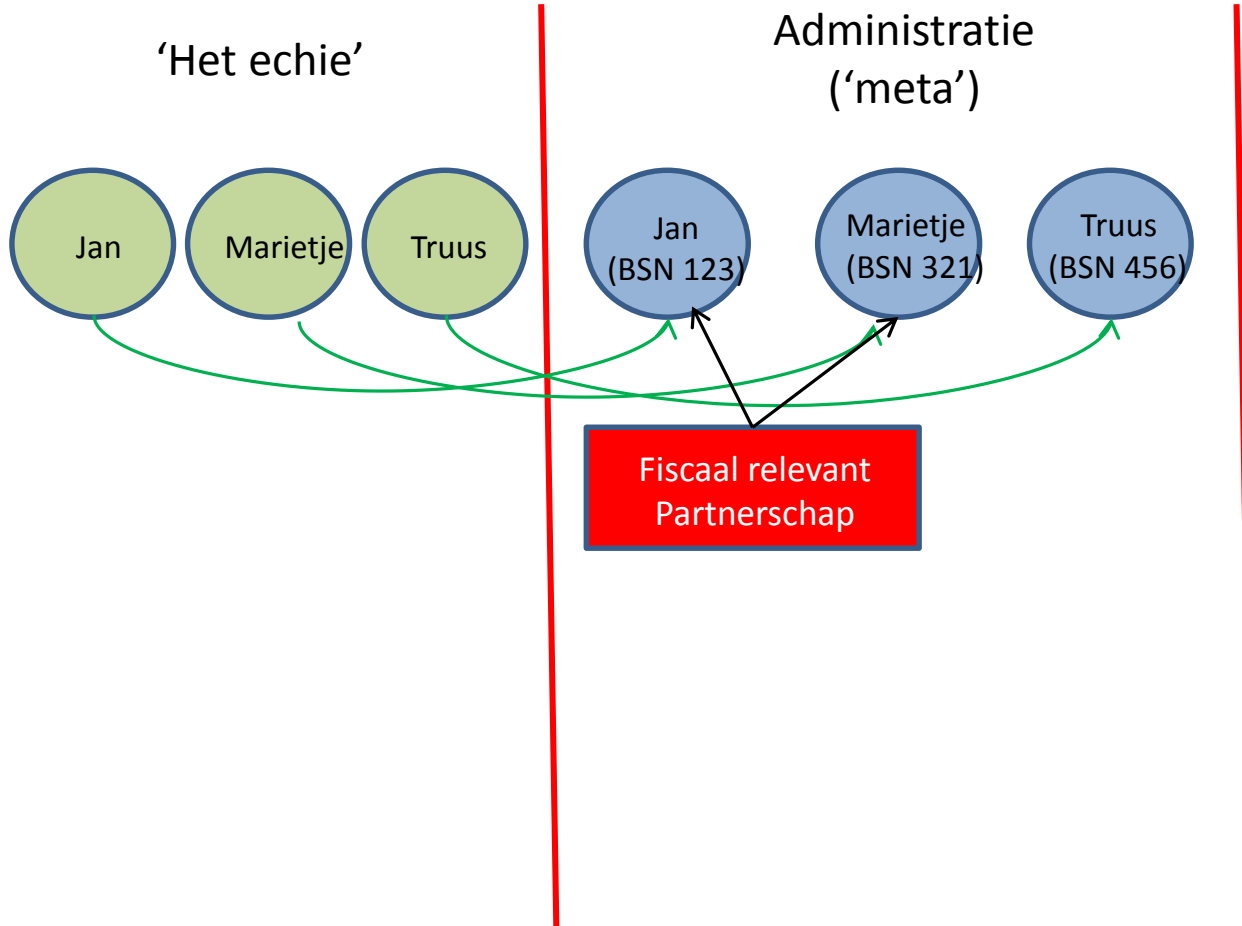


# PROV en abstractieniveaus



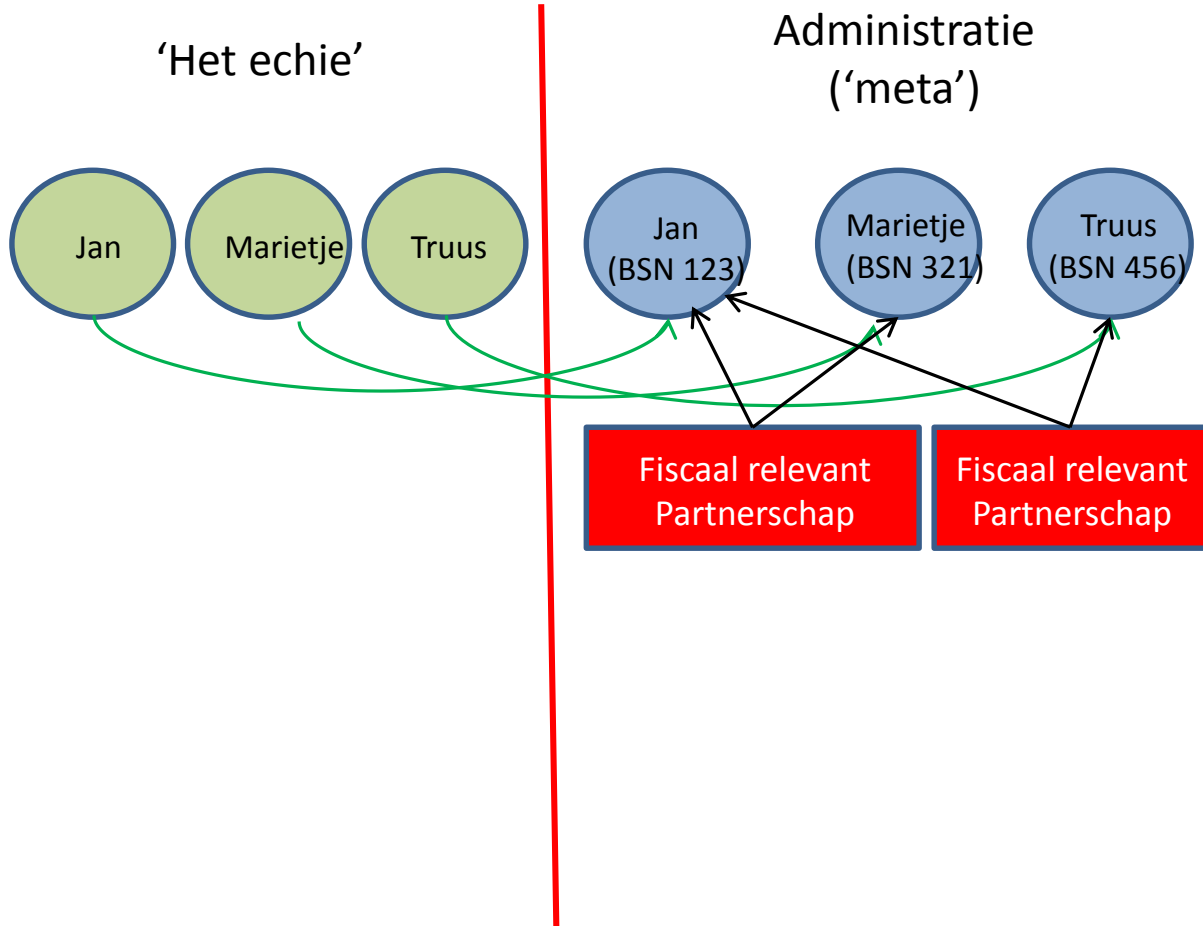
Procesverloop  
administratie  
(‘metameta’)

# PROV en abstractieniveaus



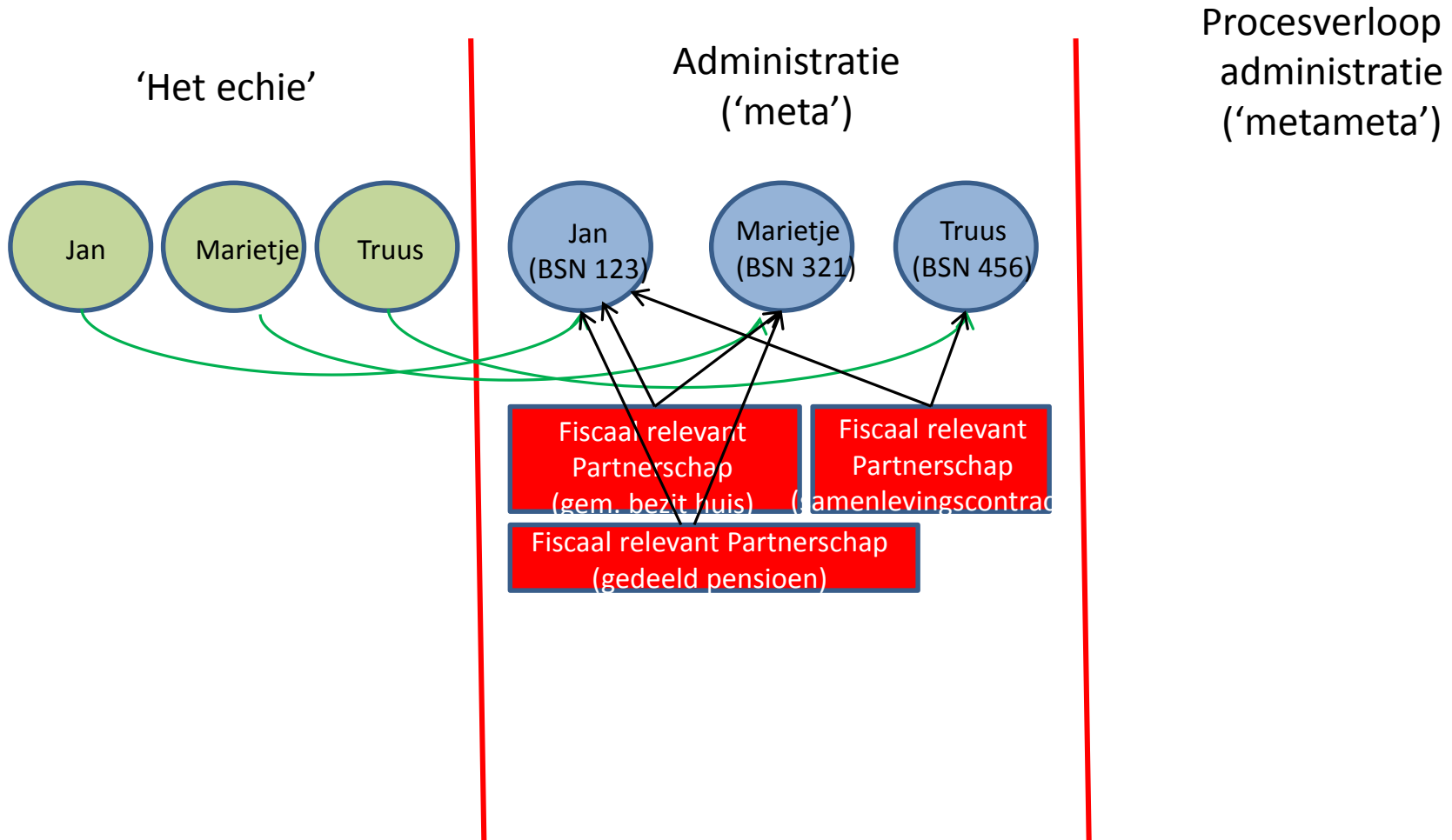
Procesverloop  
administratie  
(‘metameta’)

# PROV en abstractieniveaus

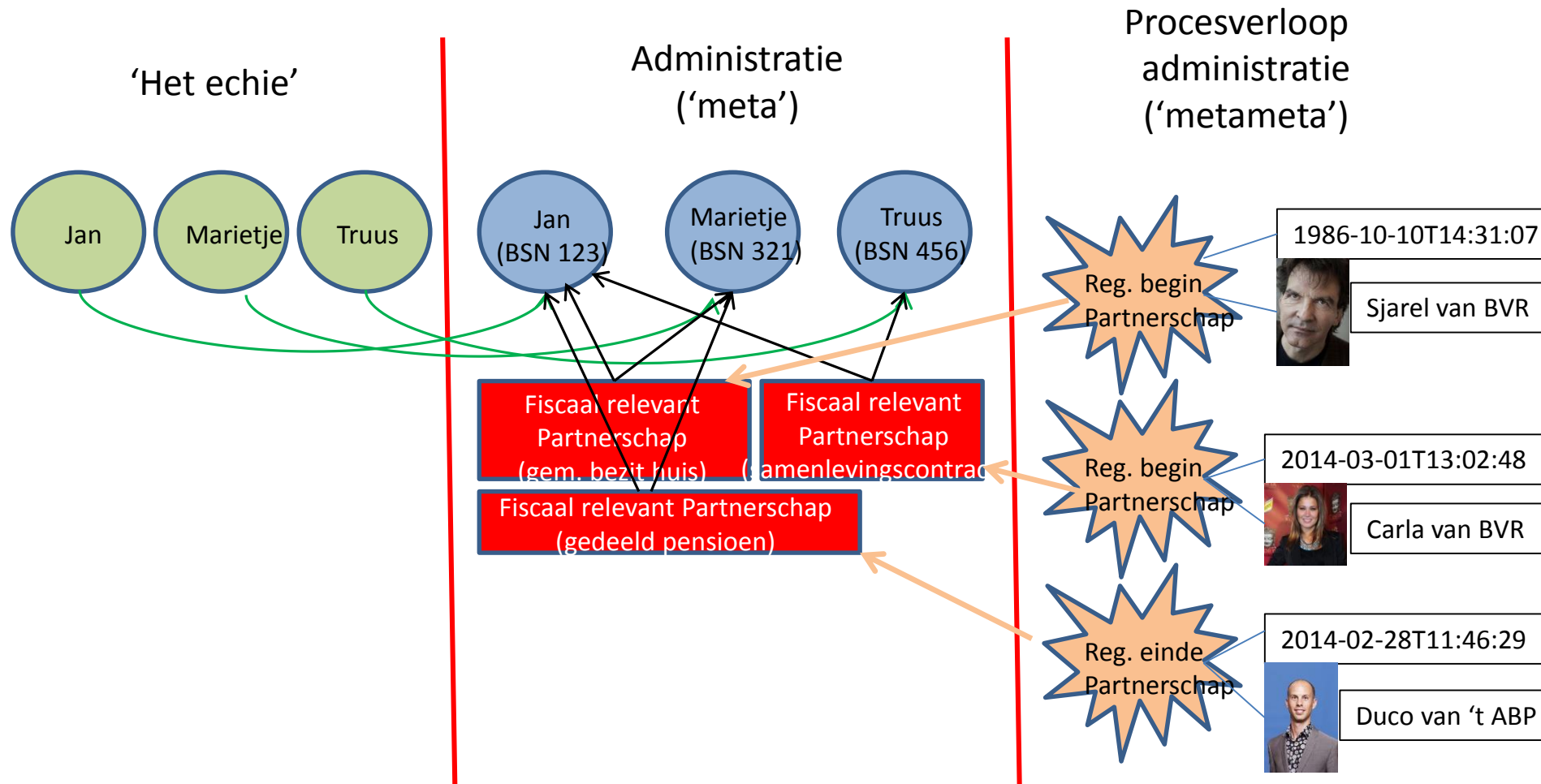


Procesverloop  
administratie  
(‘metameta’)

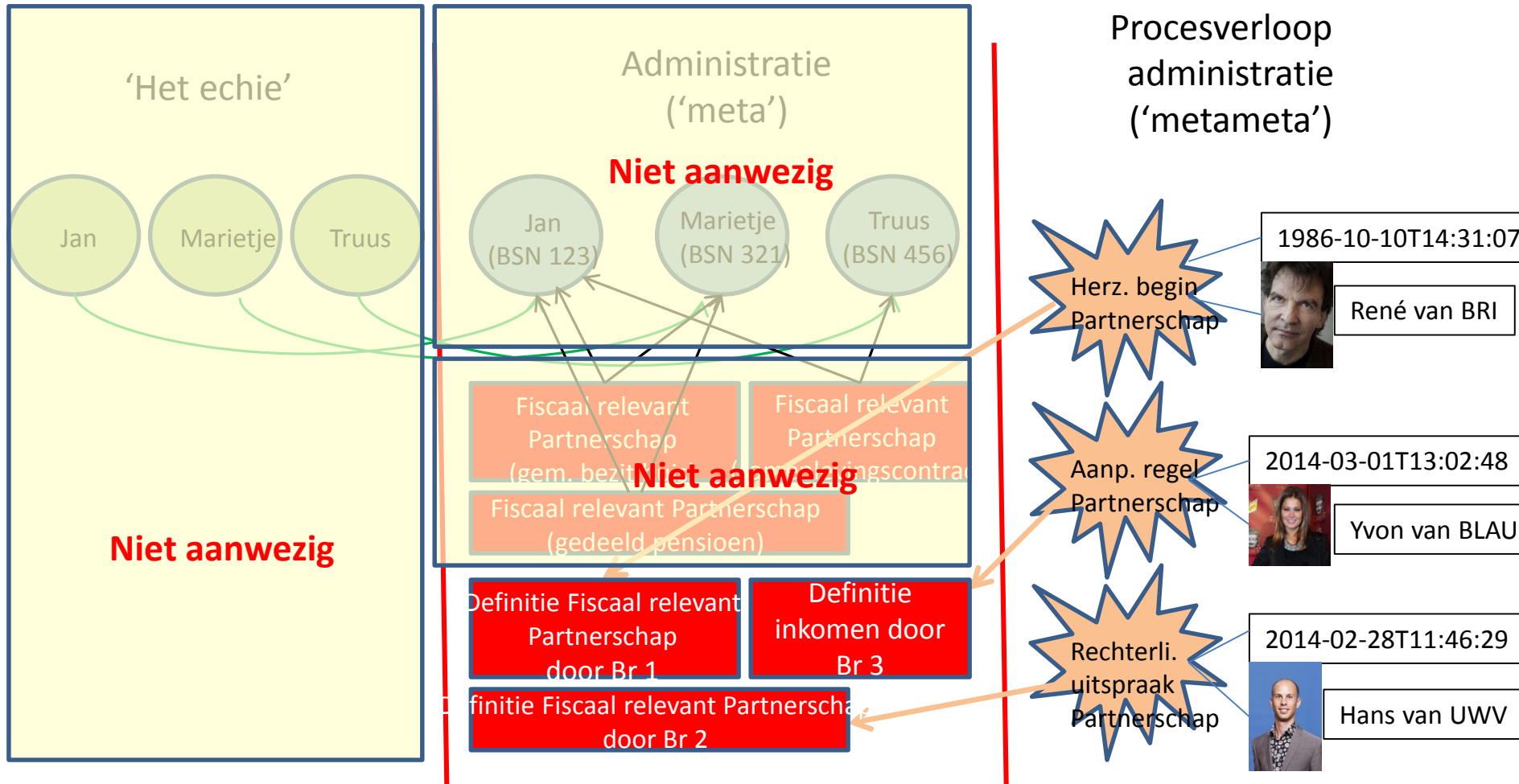
# PROV en abstractieniveaus



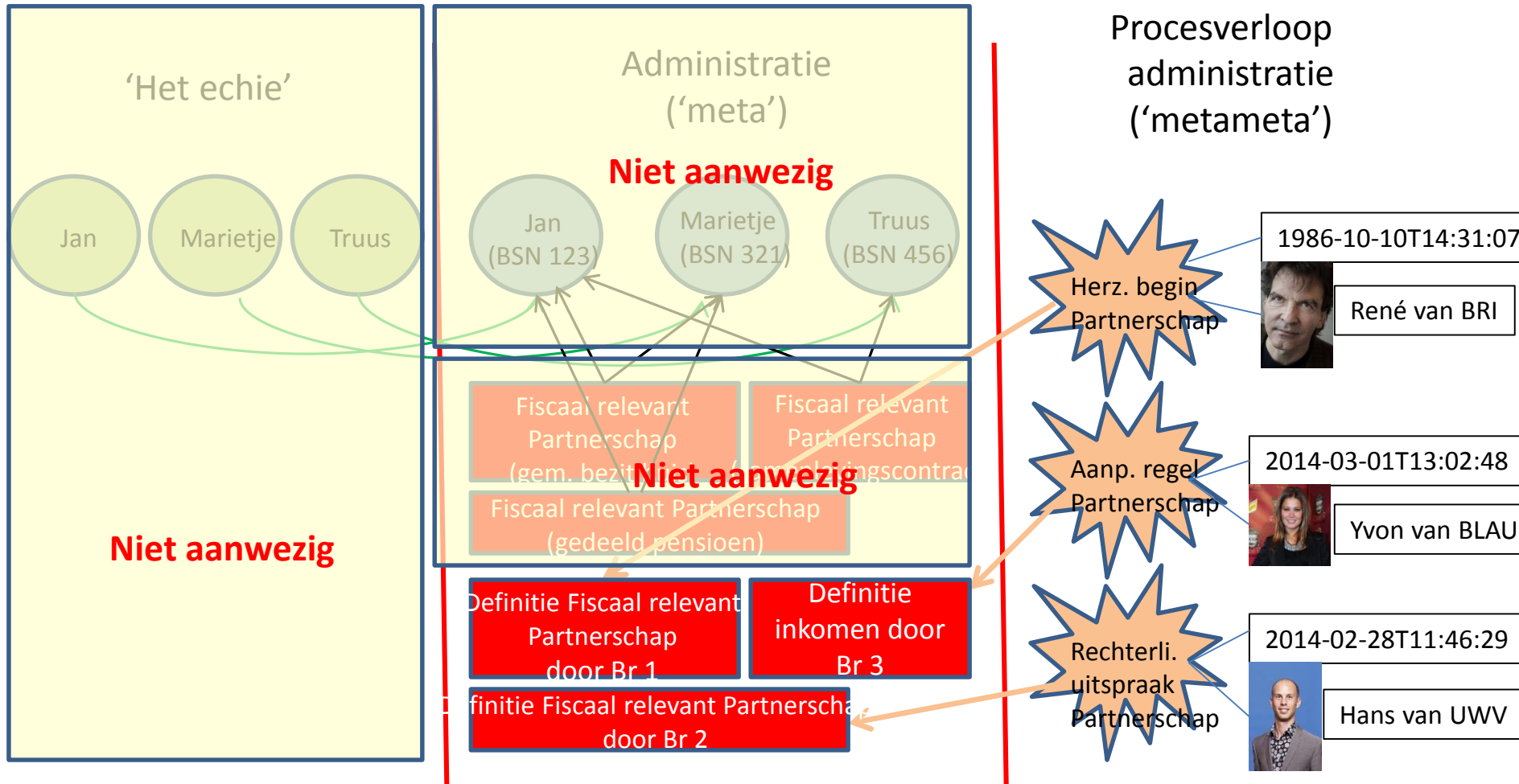
# PROV en abstractieniveaus



# PROV en de Stelselcatalogus



# PROV en de Stelselcatalogus



# Wat moet voor PROV in de Stelselcatalogus gebeuren

- Scheiding doorvoeren tussen abstract metadataveld en de versies daarvan

1. Begripvoorkomens, ~~GegevensElementen~~ en Begrippen worden opgevat als versieloze placeholders, die abstracte Begripvoorkomens enz. koppelen aan concrete versies. De individuele Begripvoorkomens enz. behouden hun huidige URI, maar worden additioneel getypeerd als zijnde van de klassen ~~prov:Entity~~ en ~~sioc:Post~~. De gedachte achter het

- Met spoed toewerken naar een referentielijst voor overheidsorganisaties (à la oog: vocabulaire van KOOP, liefst ook mede op basis van OIN)
- Inbouwen van procedures (in user interface, validaties in Sesame of TopBraid Composer) die PROV-ketens genereren, *bundles* met provenance-entiteiten
- Verantwoordelijkheden voor versies beleggen



# PROV entiteiten in de Stelselcatalogus

Klasse: :Basisregistratie  
Superklasse: :Registratie

Basis-  
registratie

dcterms:isPartOf

dcterms:isPartOf

:basisregistratie

Klasse: skos:ConceptScheme

Verzameling  
van  
begrippen

:basisregistratie

'Begrip' en 'BegripVoorkomen'  
zijn subklasse van skos:Concept

skos:inScheme

Begrip

Begrip-  
Voorkomen

Gegevens-  
Element

Wet- en  
Regelgeving

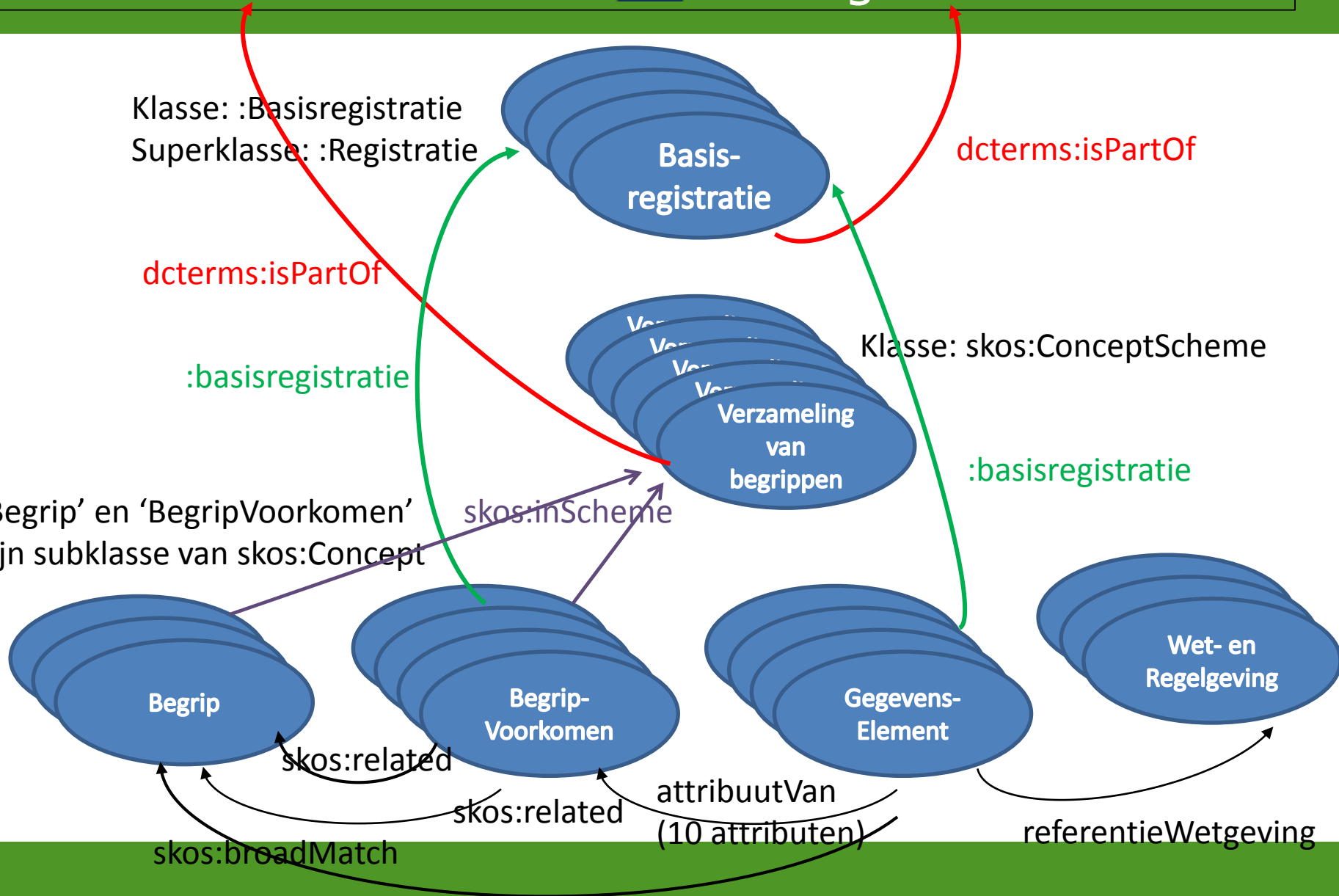
skos:related

skos:related

attribuutVan  
(10 attributen)

referentieWetgeving

skos:broadMatch





Vragen?