

Kennissessie Metadata

Onderdeel van het NORA Open Huis van de Architectuur

&

Voorprogramma bij Lexi Rowland's verdediging
van haar Engineering Doctorate

Programma & Sprekers

13.30 Welkom & Introductie:

- **Arjen Santema** (Kadaster, voorzitter NORA Expertgroep Gegevensmanagement)
- **André van Brussel** (ICTU, FDS)
- **Marco Aarts** (ICTU, Werkgroep Metadata Standaarden)
- **Danny Greefhorst** (BZK, bureau MIDO)
- **Lexi Rowland** & begeleider **Erwin Folmer** (Kadaster, Universiteit Twente)

13.45 – 14.30 Presentaties

- **Nationaal Semantisch Vlak**
- **Metadata in het Federatief Datastelsel**
- **Naar een Stelsel van Metadatastandaarden**
- **Metadata voor gegevensuitwisseling**

14.30 (ICTU+online): Werksessie Principes voor metadata voor gegevensuitwisseling

14.30 (UT): Verdediging Lexi Rowland

kadaster



Arjen Santema | 5 maart 2024

Gegevensmanagement nationaal semantisch vlak



Fysieke werkelijkheid – juridische werkelijkheid

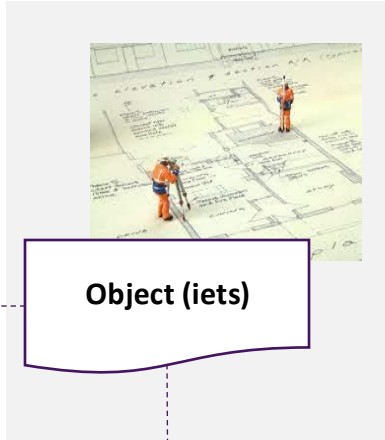
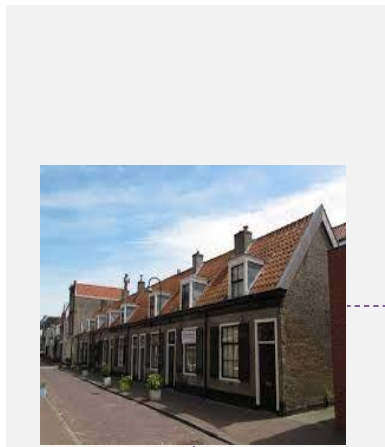
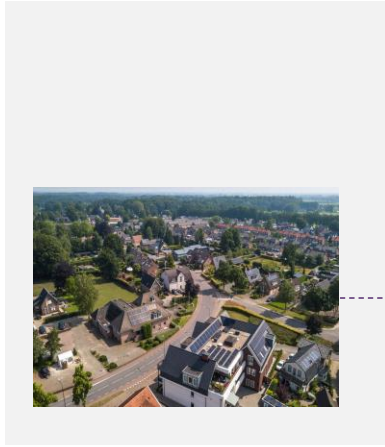


leefomgeving

juridische werkelijkheid

leefomgeving

juridische werkelijkheid



Wet- en
regelgeving

Object (iets)



gebeurtenis



Gegevensmanagement – nationaal semantisch vlak



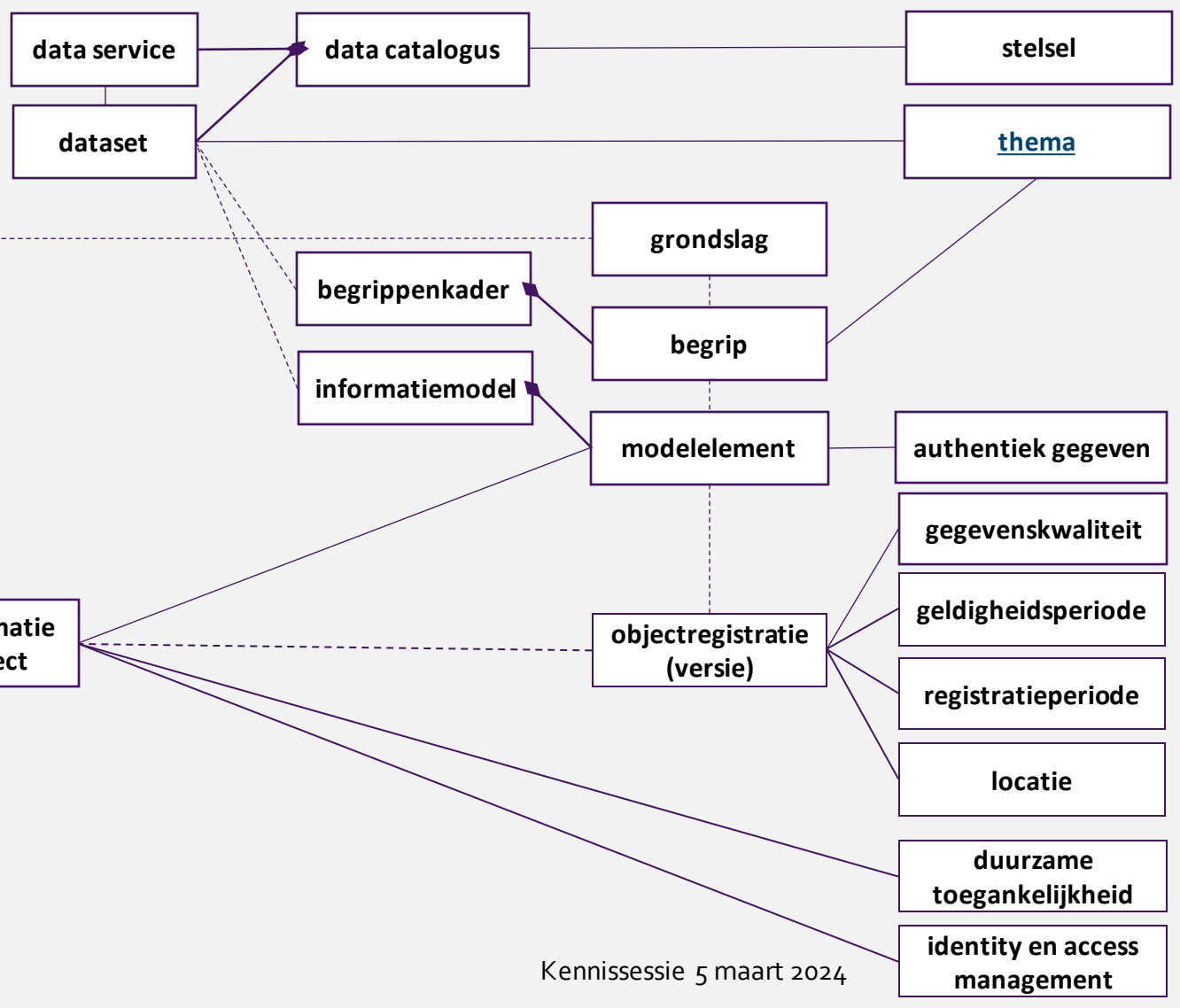
Juridische werkelijkheid

administratieve werkelijkheid: data

Metadata

Standaarden

Vocabulaires



wet- en regelgeving



object (iets)



gebeurtenis

—◆ is opgenomen in
— semantische samenhang
- - - - - samenhang

NL api-strategie

DCAT-AP-NL

DCAT

BWB

JCDR

SBB

SKOS

MIM

core vocabularies

SHACL

RDFS

OWL

ISO 25012

ISO 25024

DQV

NEN 2660

NEN 3610

DC-terms

vc

prov-o

owl-time

MDTO

E-Idas

IRMA/Yivi

SOLID



Metadata in het Federatief Datastelsel

Open Huis van de Architectuur, 5 maart 2024

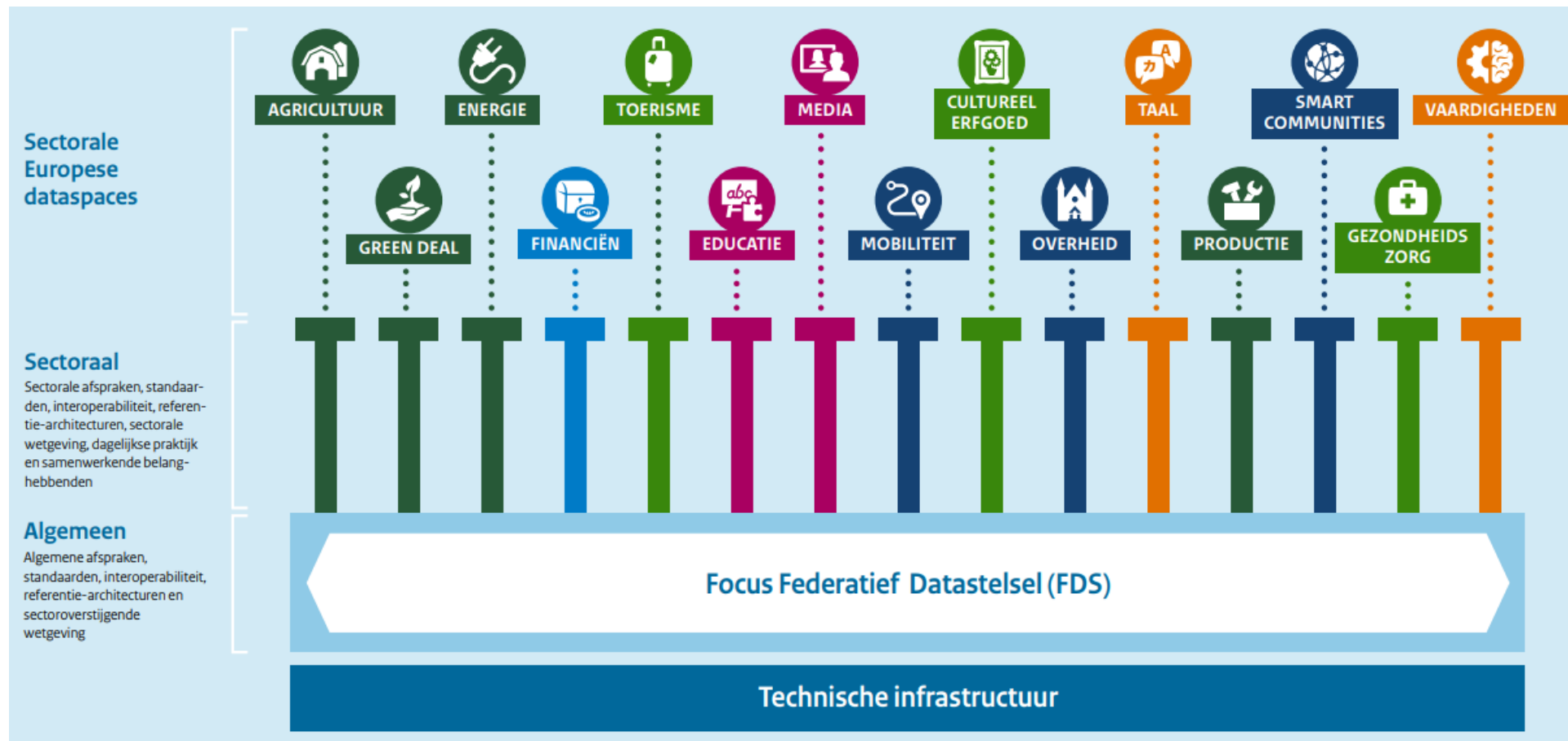
Andre van Brussel

ICTU

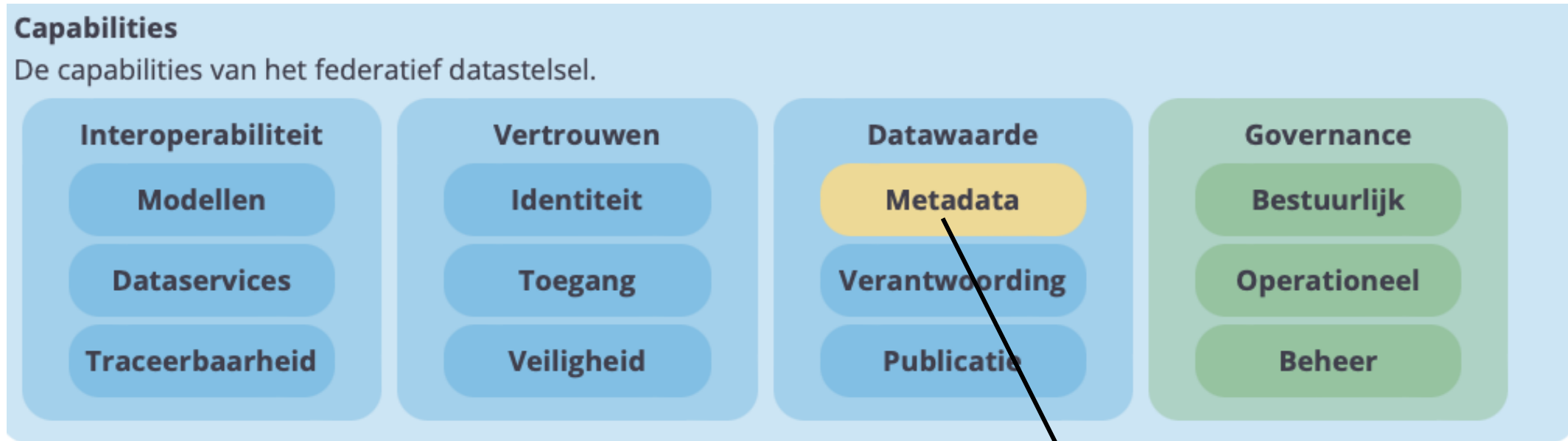
Kennissessie 5 maart 2024

Federatief Datastelsel: wat is het

Cross-sectoraal datadelen op verantwoorde wijze



Federatief Datastelsel: wat moet het kunnen

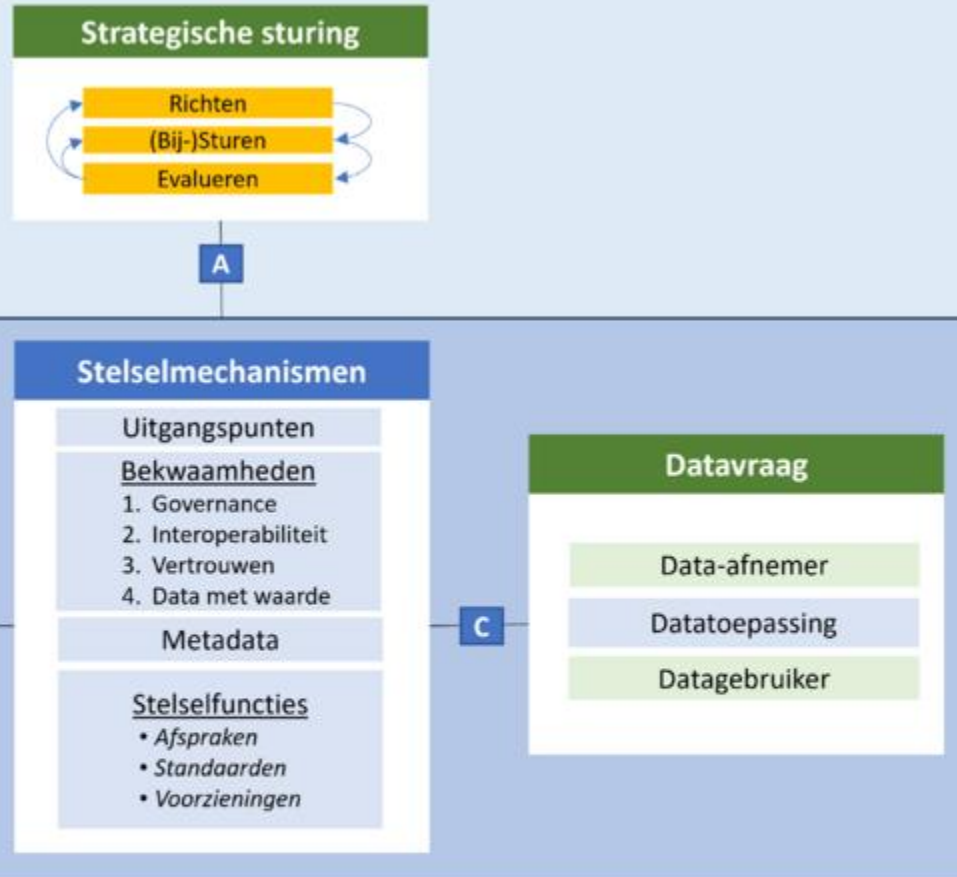


Ontleend aan OpenDEI

Gemeenschappelijke (zelf)beschrijvingen met verwijzingen:

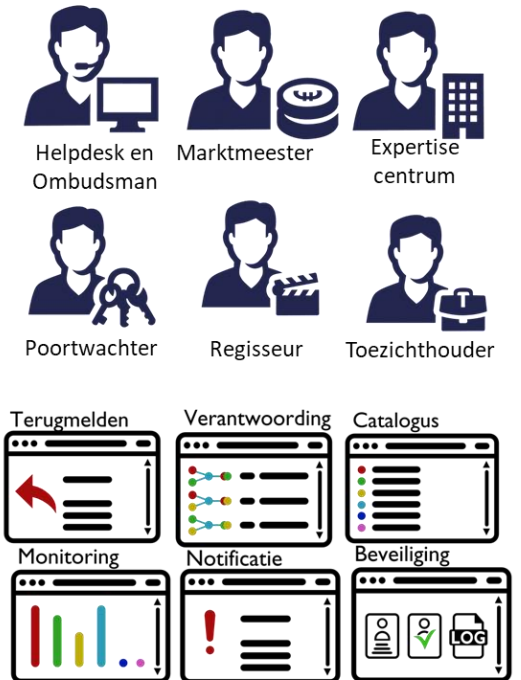
- Bronnen
- Diensten
- Deelnemers.

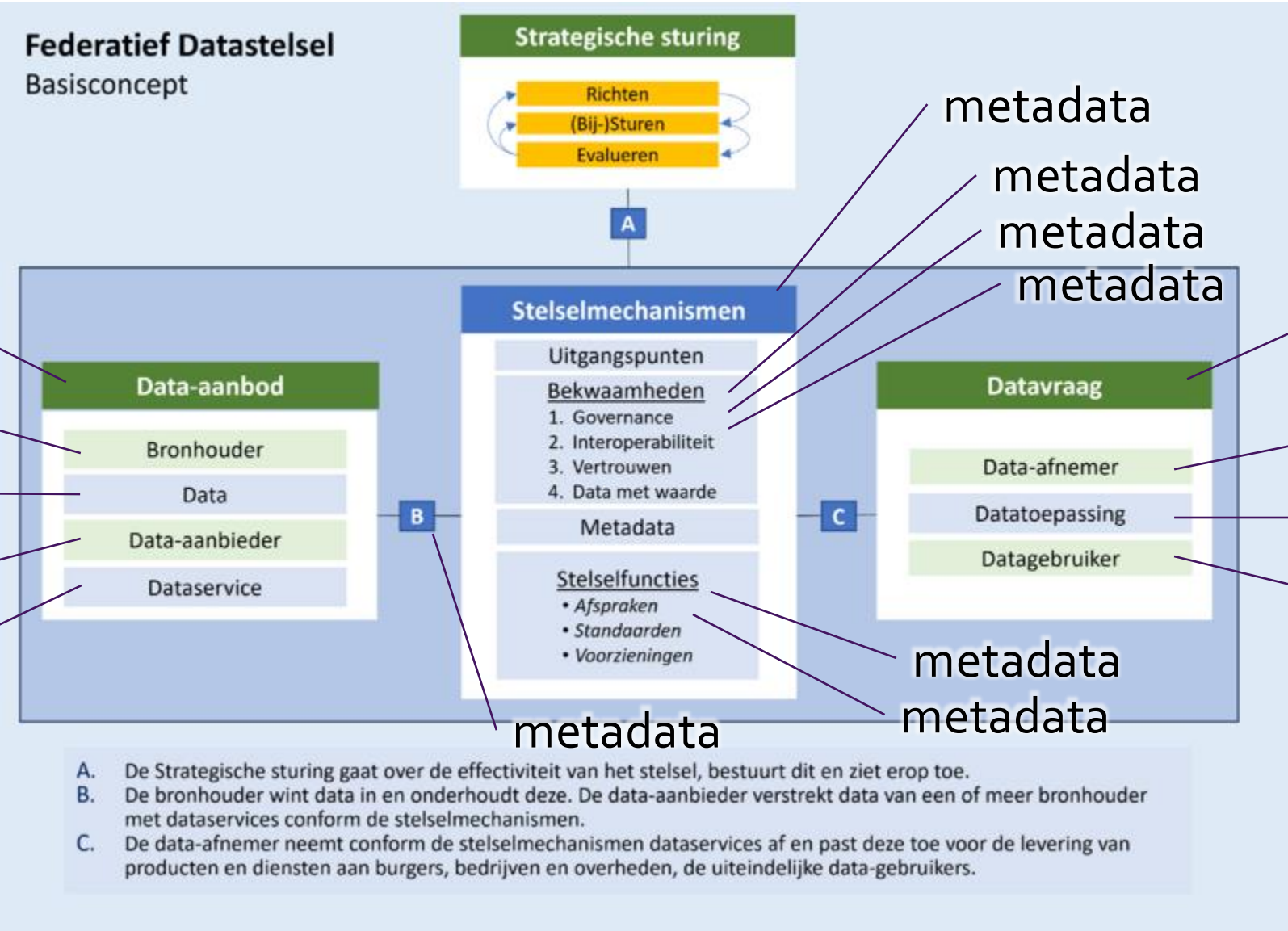
Federatief Datastelsel Basisconcept

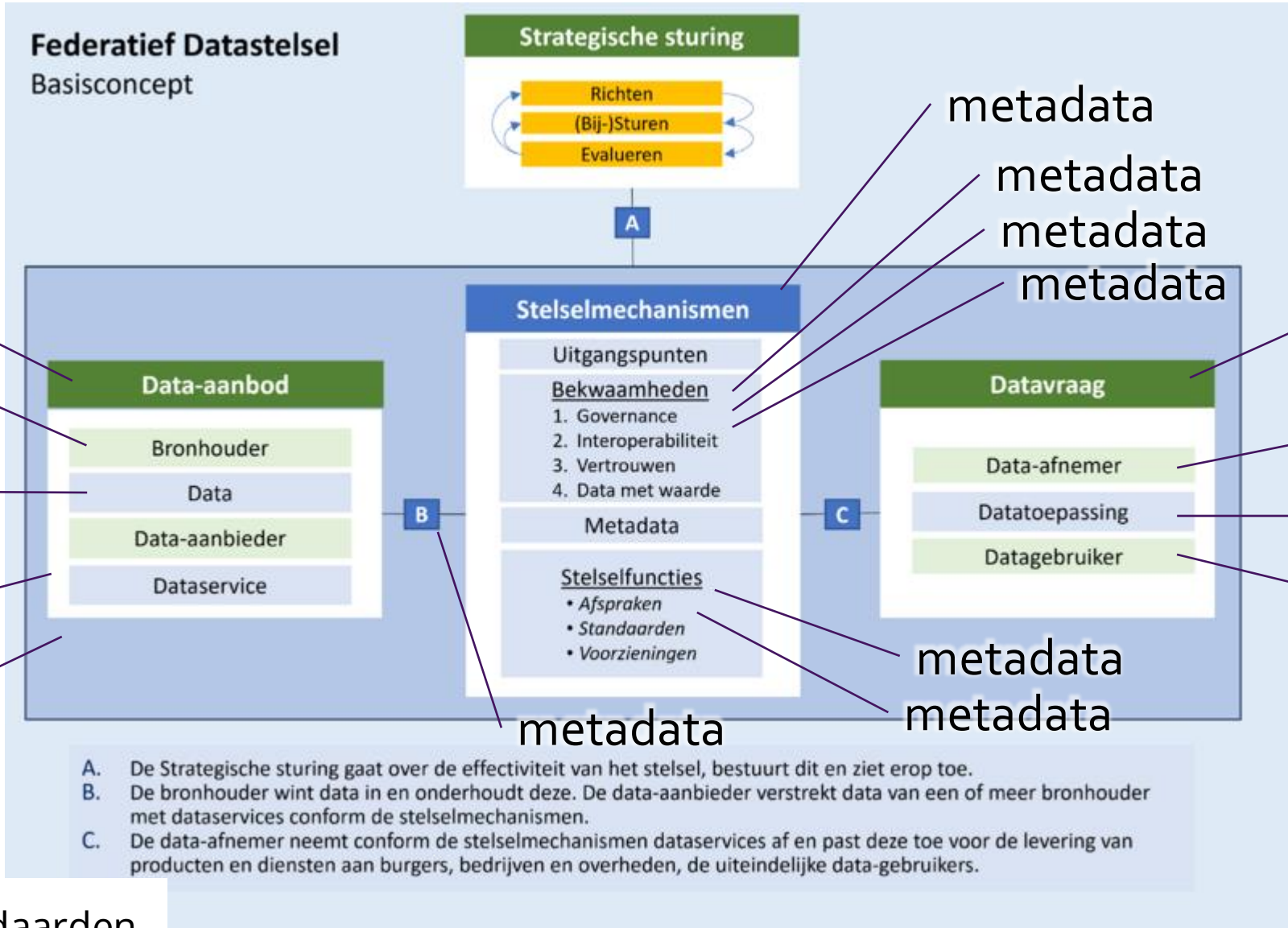


- A. De Strategische sturing gaat over de effectiviteit van het stelsel, bestuurt dit en ziet erop toe.
- B. De bronhouder wint data in en onderhoudt deze. De data-aanbieder verstrekt data van een of meer bronhouder met datadives conform de stelselmechanismen.
- C. De data-afnemer neemt conform de stelselmechanismen datadives af en past deze toe voor de levering van producten en diensten aan burgers, bedrijven en overheden, de uiteindelijke data-gebruikers.

De stelselfuncties







metadata

metadata

metadata

Datasets, Begrippen, Informatiemodellen

metadata

metadata

Dataservices

Standaarden geselecteerd

metadata

metadata

metadata

metadata

metadata

metadata

metadata

metadata

metadata

metadata

metadata

Aan de slag?

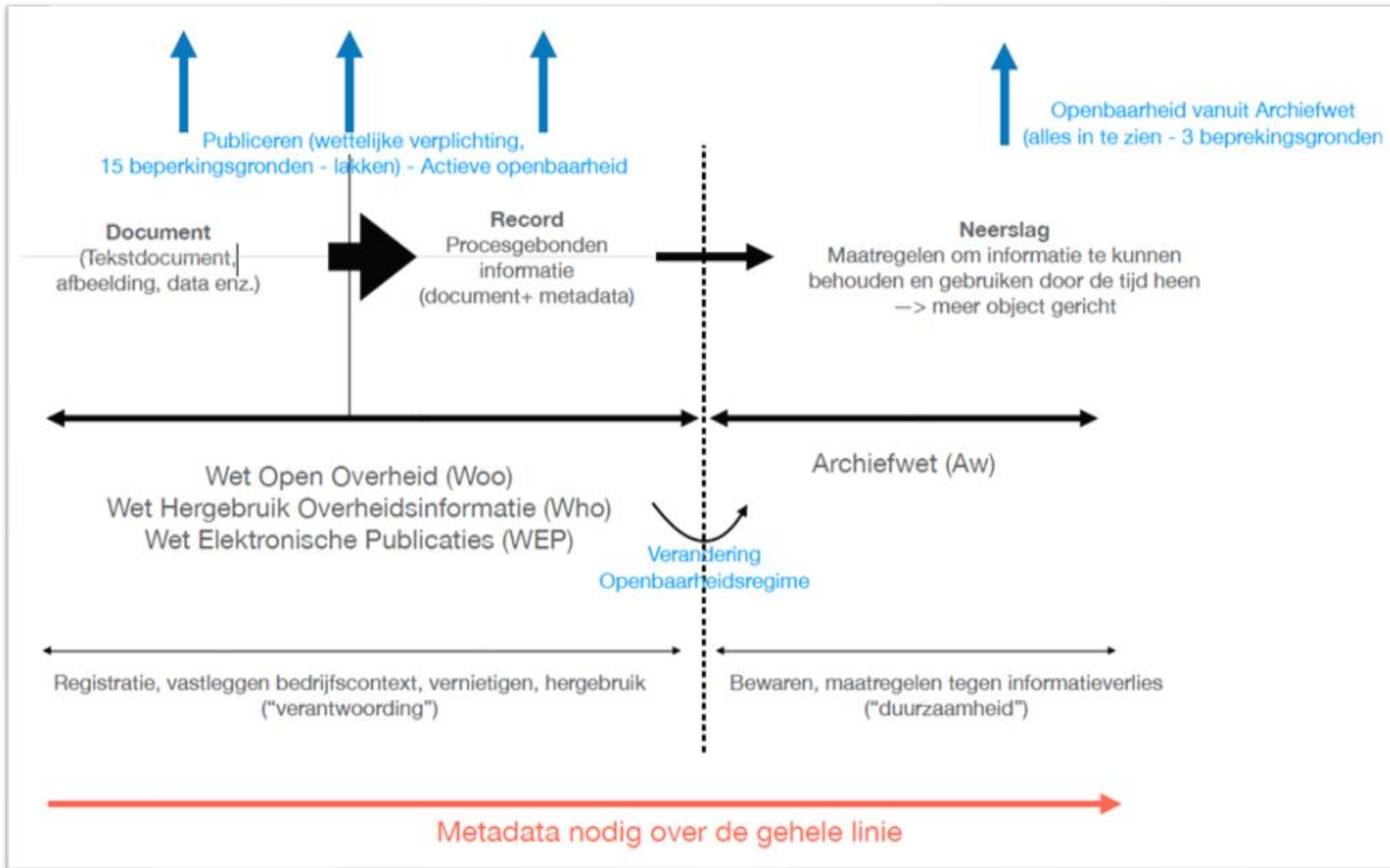
- Meehelpen met de metadatapuzzels:
<https://federatief.datastelsel.nl/docs/realisatie/capabilities/datawaarde/metadata/>
- Beschrijf je data-aanbod volgens de standaarden:
Datasets, dataservices, begrippen en informatiemodellen
- Algemene informatie IBDS en FDS: <https://realisatieibds.pleio.nl>

Naar een Stelsel van Metadata Standaarden

Werkgroep Metadata Overheidsinformatie in 2023 en 2024

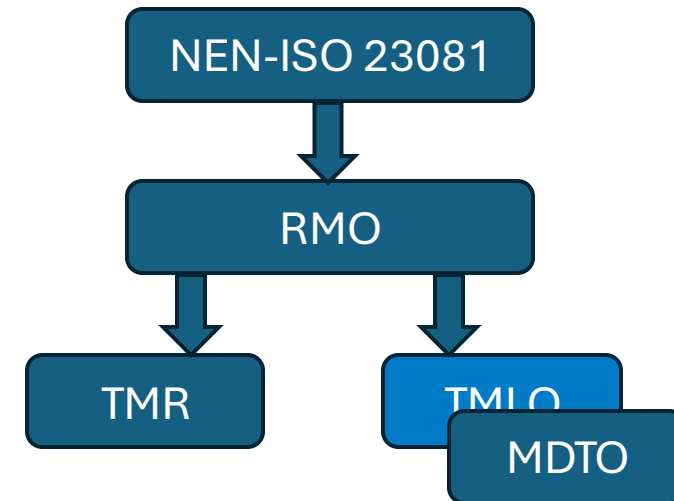
Werkgroep metadata overheidsinformatie

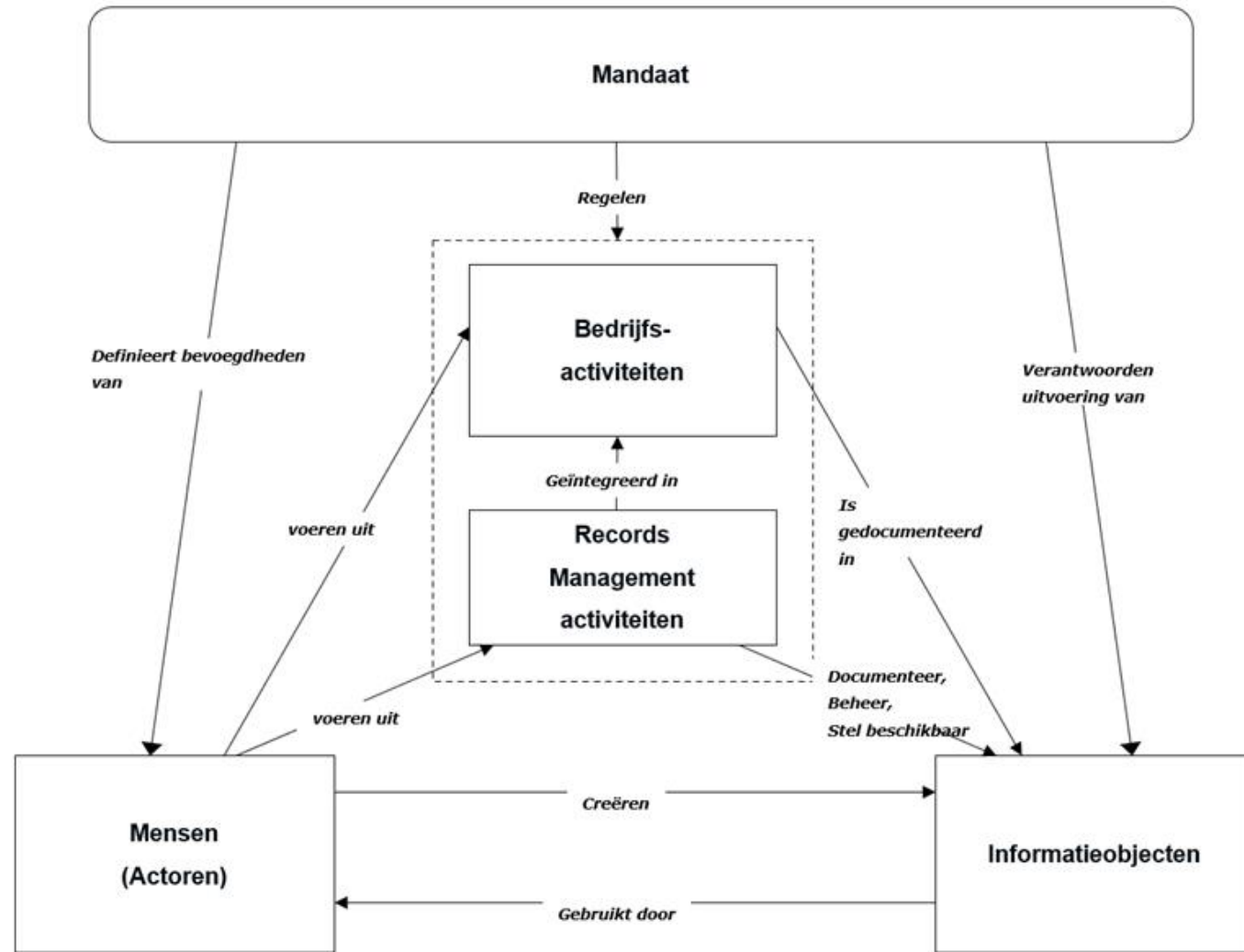
- Onderdeel van Rijksprogramma voor Duurzaam Digitale Informatiehuishouding RDDI
- Informatie- en archiefbeheer
- Uitwisseling en overbrenging
- Werkgroep Metadata Overheidsinformatie richt zich vooral op het **creëren en besturen van een stelsel van metadata standaarden**
- Vertegenwoordigers van
 - BZK/CIO Rijk
 - Nationaal Archief
 - Ministerie van Onderwijs, Cultuur en Wetenschap
 - Ministerie van Volksgezondheid, Welzijn en Sport
 - Ministerie van Veiligheid en Justitie (CIO en Justid)
 - Ministerie van Infrastructuur en Waterstaat / Rijkswaterstaat
 - Ondersteund door ICTU
 - KOOP
 - Decentrale overheden (vanaf medio 2023)



Werkgroep in 2023: Verkenning

- Onderzoek actualisering **Richtlijn Metagegevens Overheidsinformatie** (RMO)
- Onderzoek:
 - ‘gap’ **MDTO**
 - Delta versies NEN-ISO 23081¹
- Uitvraag bij grote programma’s (WaU, OoO, FDS, RoG, etc.)
- Governance (stelsel + RMO)
- Conclusie: behoefte aan **overzicht** en **governance** van het geheel aan metadata standaarden





“5-entiteiten model”

- MDTO dekt ‘**informatieobject**’ af (ook wel archiefstuk of document genoemd in wetgeving)
- Standardisatie van de andere entiteiten als **bouwblok** in 2024

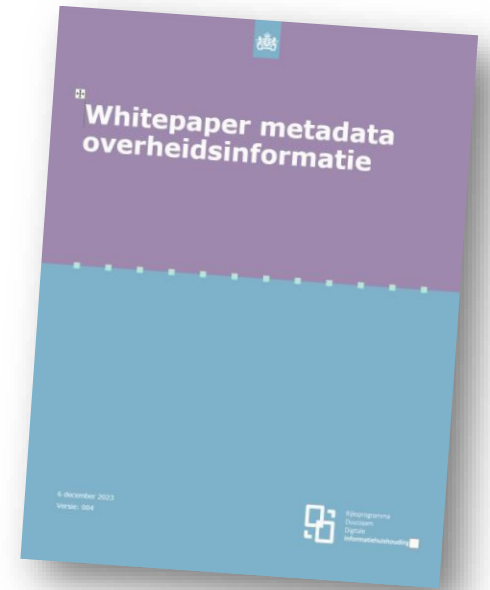
Werkgroep in 2023: Whitepaper en adviesnota

- **Whitepaper**

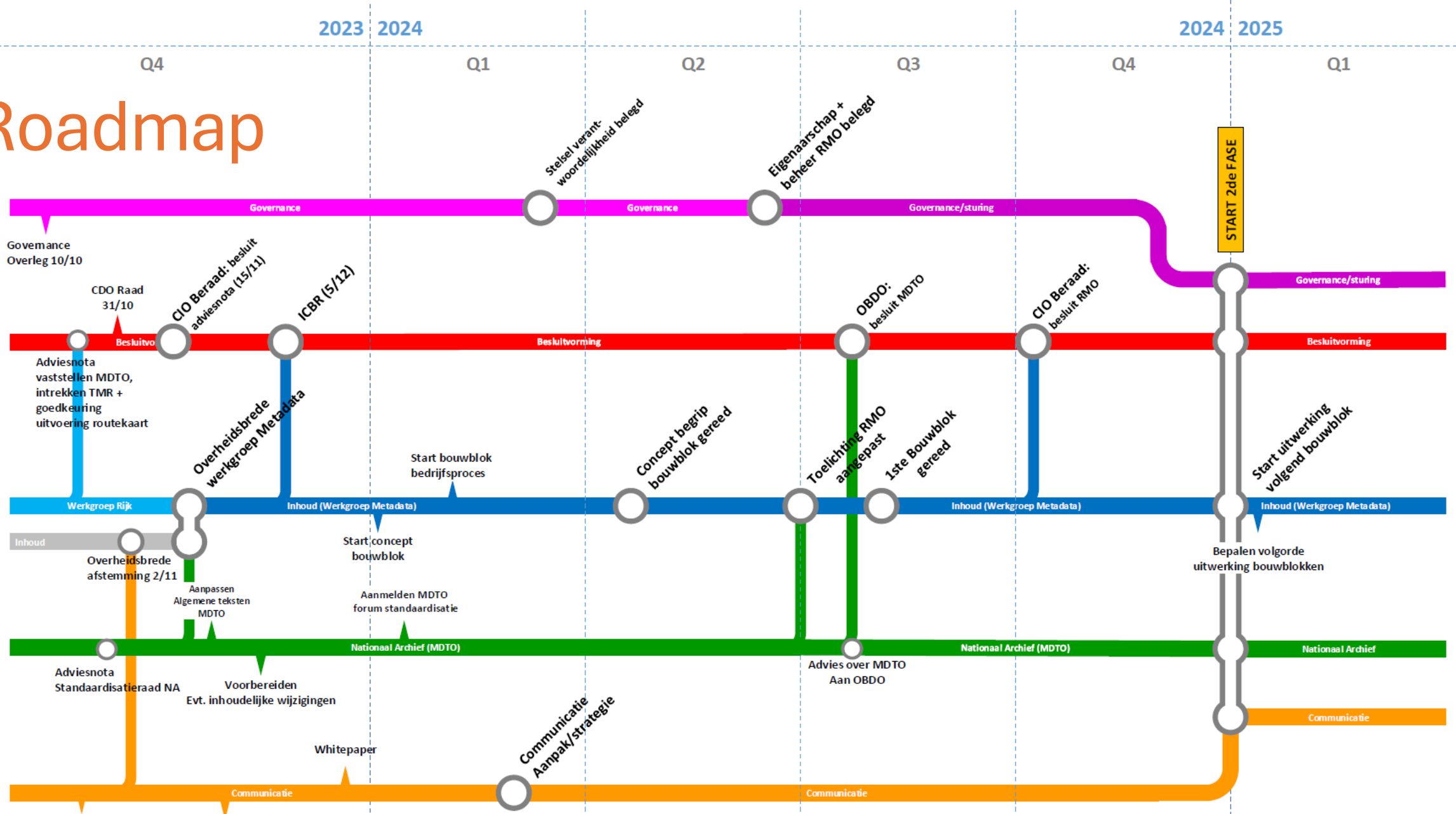
- Aanpak conform BOMOS

- **Adviesnota en Roadmap** voor CIO-beraad en ICBR

- MDTO omarmd, TMR ingetrokken
- Akkoord met uitvoering roadmap



Roadmap



Kennissessie 5 maart 2024

naar een stelsel van metadata standaarden

Belangrijkste activiteiten 2024

- Aanmelding MDTO bij **forum standaardisatie**
- **Governance**
 - Stelsel: eigenaarschap en beheer
 - Stelselniveau
 - individuele standaarden
 - Eigenaarschap en beheer van RMO
- **Bouwblokken**
 - Actualisatie RMO – beschrijving van de samenhang
 - Bedrijfsactiviteiten
 - Actoren
 - Mandaten



Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties

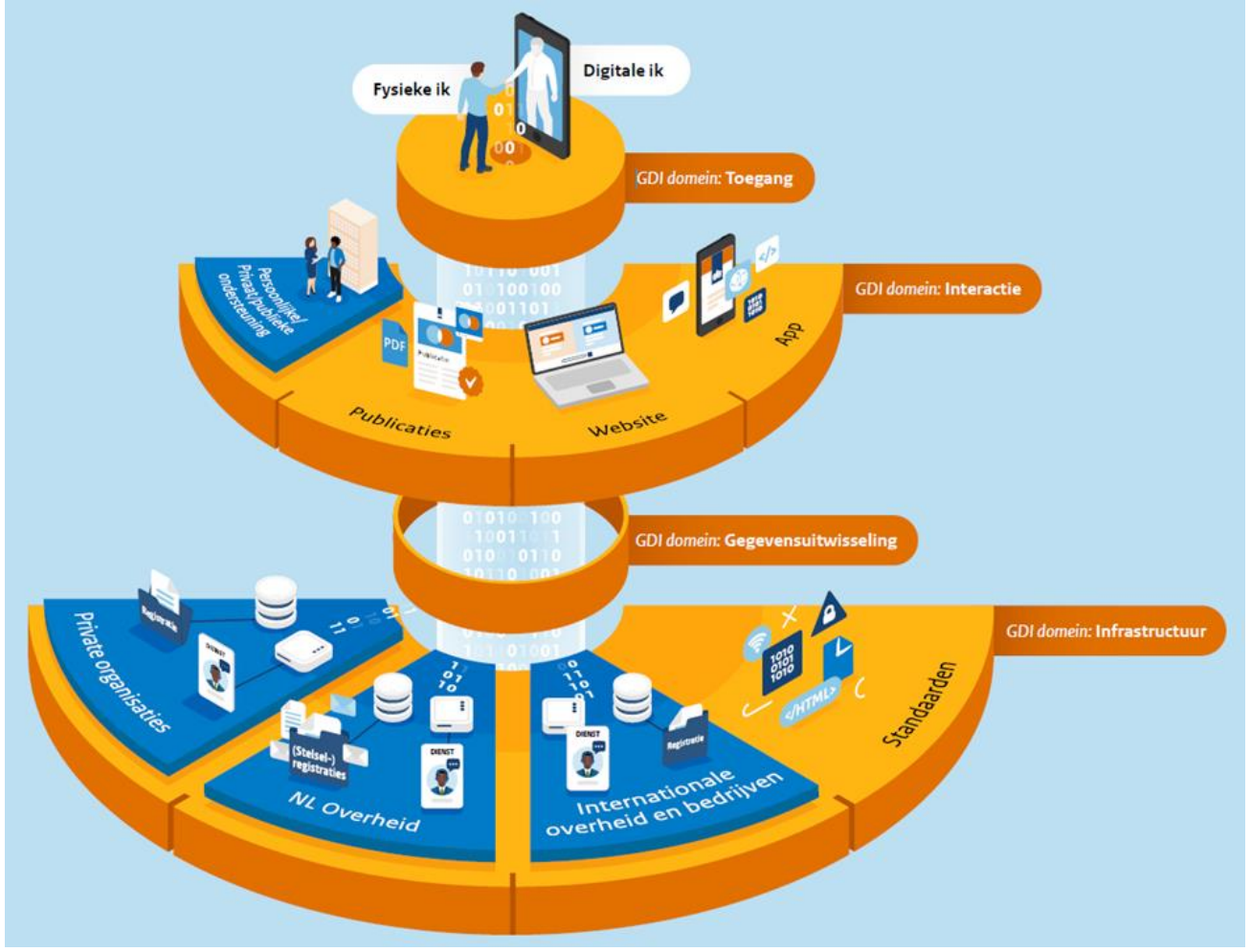


Metadata voor gegevensuitwisseling

Danny Greefhorst (MinBZK/DO/Bureau MIDO)
Danny.Greefhorst@minbzk.nl



GDI domeinen





Soorten metadata

Business metadata <i>Inhoud en governance</i>	Technische metadata <i>Details, systemen, processen</i>	Operationele metadata <i>Verwerking en gebruik</i>
<ul style="list-style-type: none">• Definities en beschrijvingen van datasets, tabellen en kolommen• Business rules, transformatieregels, berekeningen en afleidingen• Datamodellen• Datakwaliteitsregels en meetresultaten• Frequentie/timing waarmee de data worden geüpdatet• Herkomst• Datastandaarden• Aanwijzing bronsysteem• Toegestaan gegevensbereik• Vertrouwelijkheidsniveaus• Bekende issues rondom de data• Aanwijzingen voor datagebruik	<ul style="list-style-type: none">• Namen van fysieke tabellen en kolommen• Kolomeigenschappen• Eigenschappen van het databaseobject• Toegangsrechten• CRUD regels• Fysieke datamodellen, inclusief namen, sleutels en indices• Relaties datamodellen en fysieke assets• Details over gebruikte ETL• Bestandsformaat en schema definitie• Bron-naar-doel mappen• Data lineage• Namen en beschrijvingen van programma's en applicaties• Terugkeerpatronen van automatische updateprocessen• Herstel- en backupregels• Rechten, groepen en rollen voor toegang	<ul style="list-style-type: none">• Logbestanden van batch runs• Geschiedenis van extracten en resultaten• Afwijkingen van terugkeerpatronen• 'Audit, balance, control' meetresultaten• Error logging• Toegangspatronen, frequentie en uitvoertijd van rapportages en queries• Planning, uitvoer en status van onderhoud en patching• Backup-, retentie- en herstelvoorzieningen• SLA-eisen en voorzieningen• Omvangs- en gebruiksgegevens• Archiverings- en retentieregels, bijbehorende archieven• Opschoningscriteria• Regels en afspraken rondom datadelen• Technische rollen en verantwoordelijkheden

Bron: DMBOK

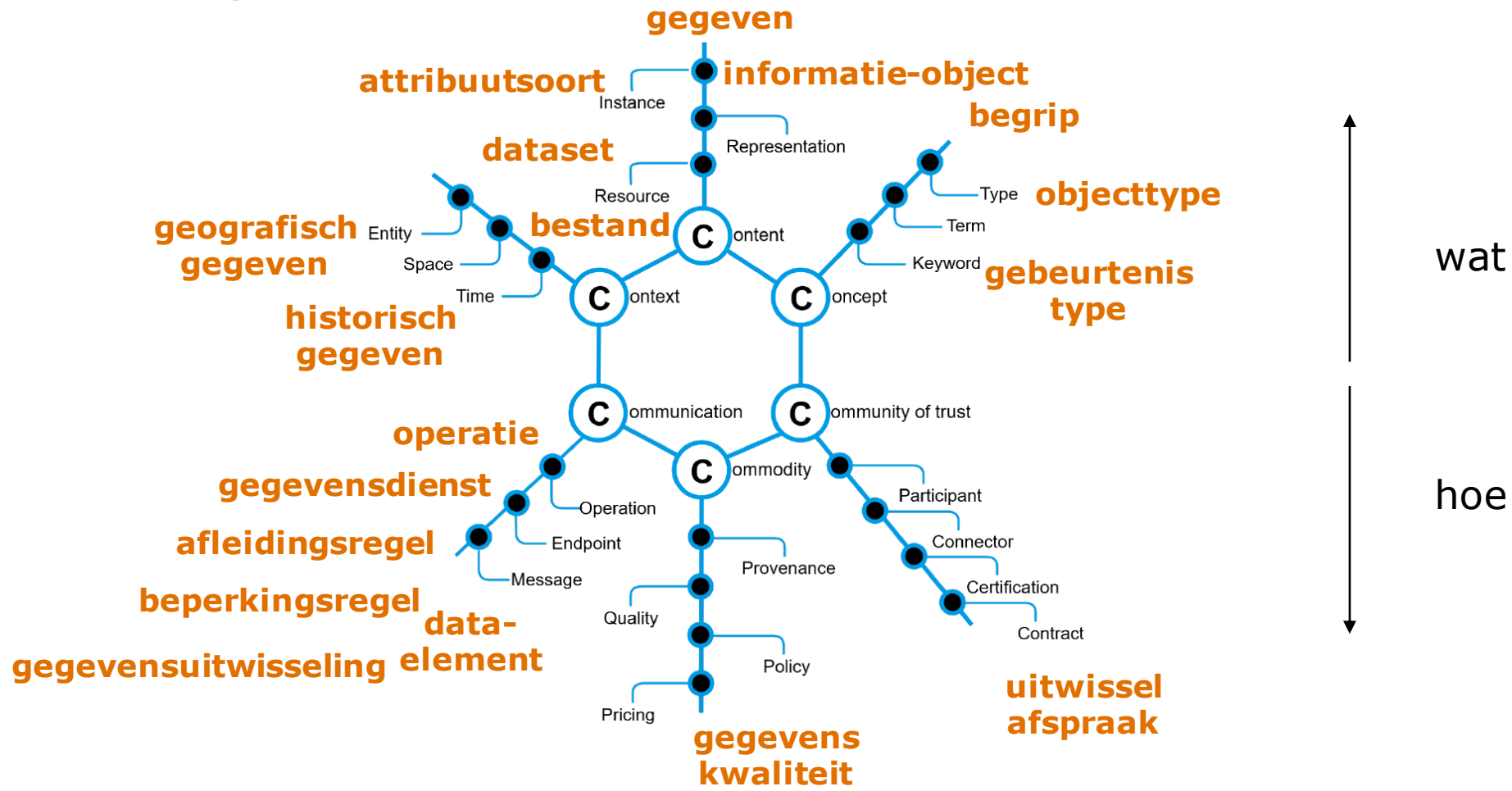


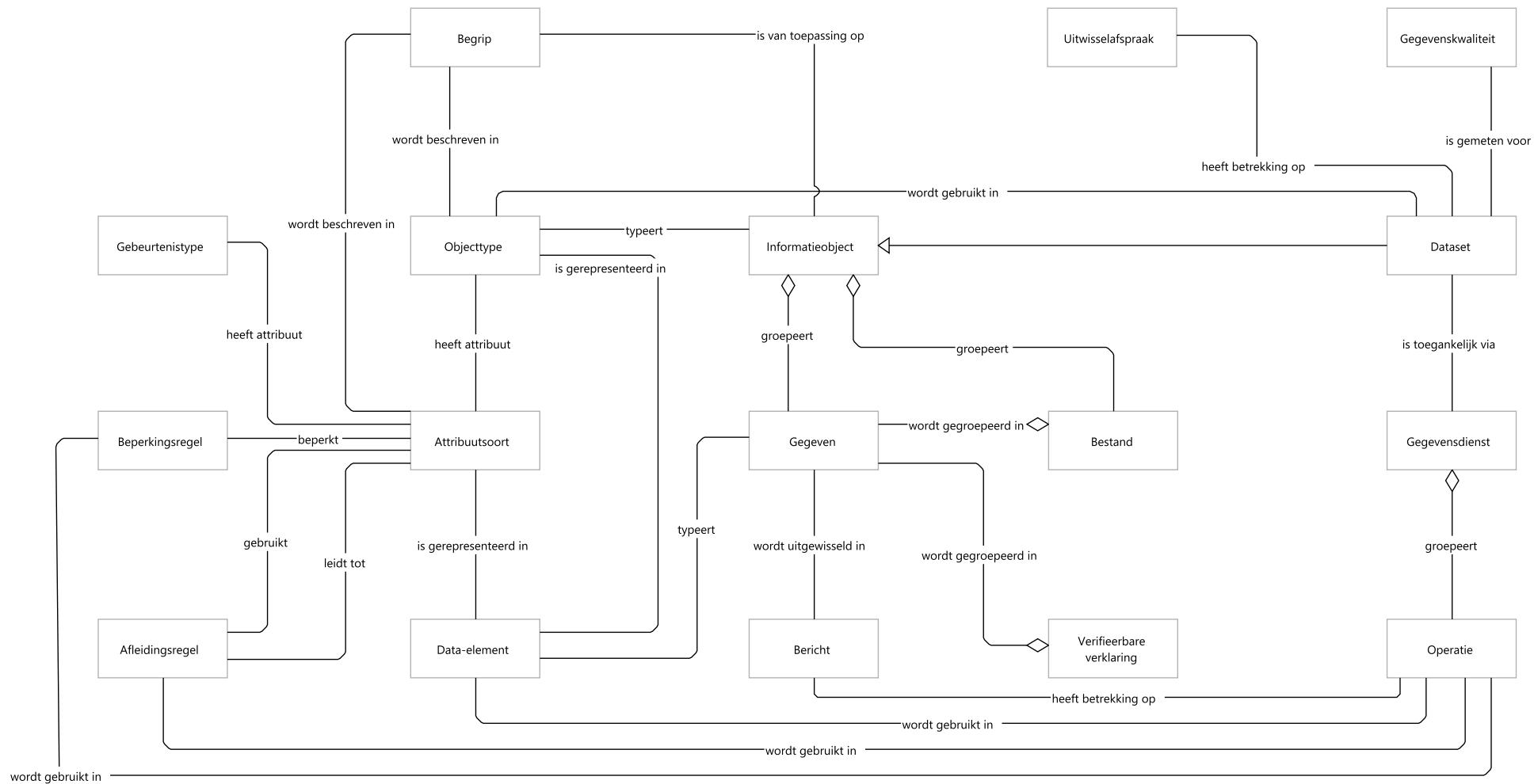
Voorbeelden van standaarden voor metadata

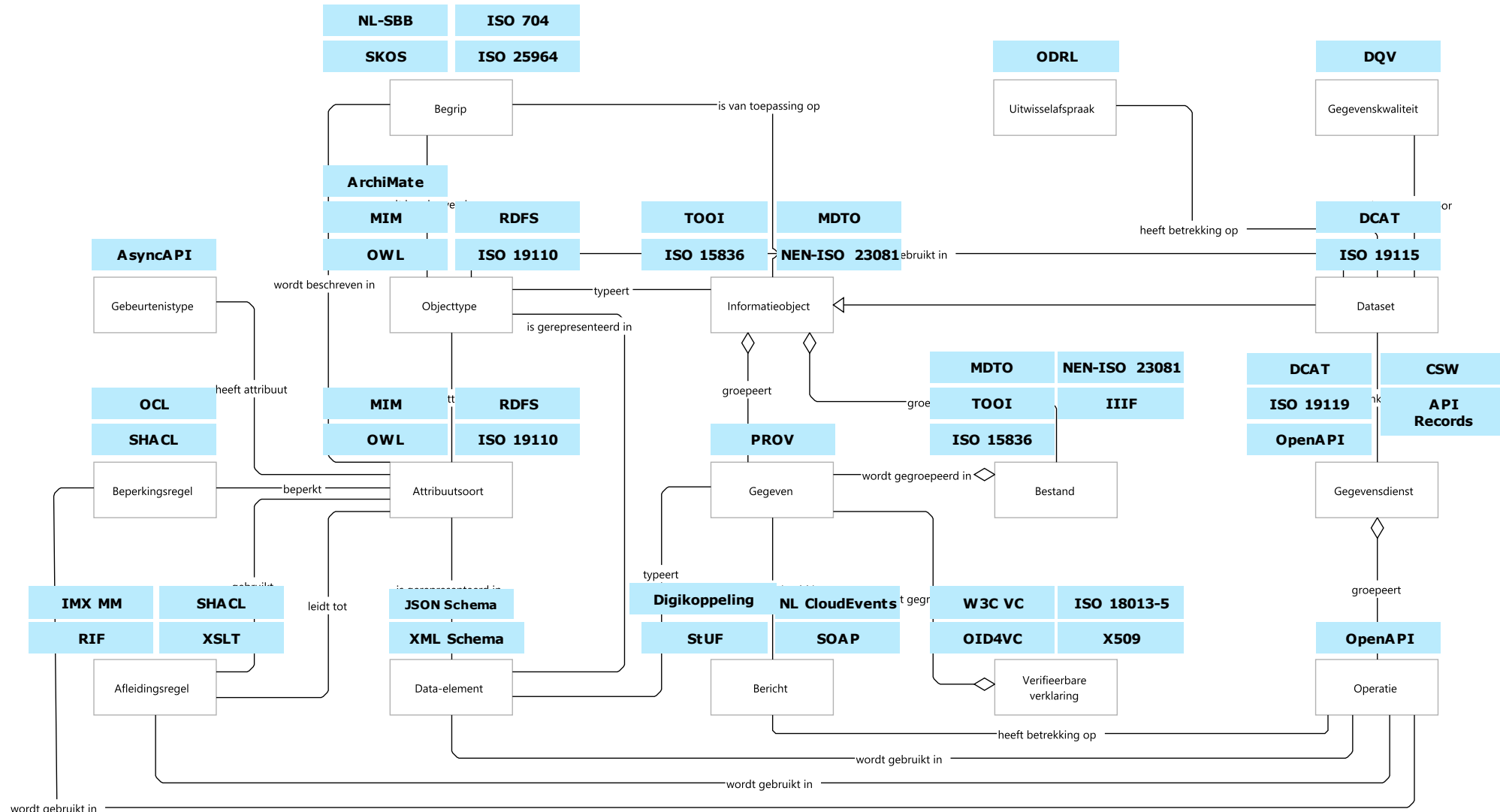
- ISO/IEC 11179
- DCAT
- ISO/FDIS 19110
- MIM
- ISO 19115
- ISO 19119
- CSW
- API Records
- ISO 25964
- ISO 704
- SKOS
- ISO 15836
- OWMS
- TOOI
- NEN-ISO 23081-2
- MDTO
- NL-SBB
- OpenAPI
- AsyncAPI
- JSON Schema
- XML Schema
- PROV
- DQV
- OCL
- OWL
- RDFS
- SHACL
- XSLT
- IMX MM
- RIF
- ODRL
- ArchiMate
- IIIF
- StUF
- Digikoppeling
- SOAP
- NL GOV profile for CloudEvents
- Verifiable Credentials
- NEN-ISO/IEC 18013-5
- OID4VC
- X509



Metadata hexagon in International Data Spaces architectuur









Meer informatie

Architectuur Digitale Overheid 2030: <https://pgdi.nl/ado>

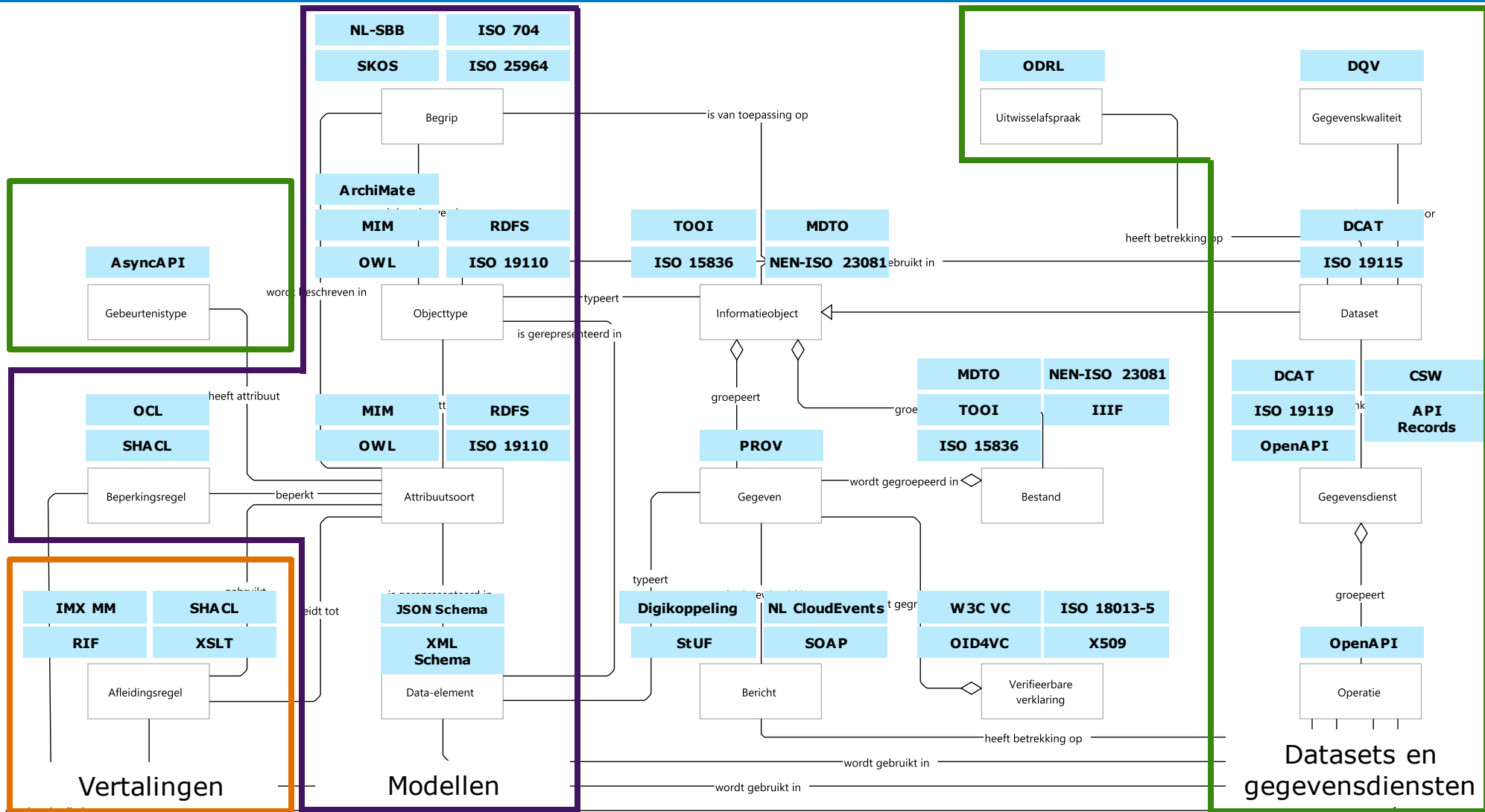
Domeinarchitectuur gegevensuitwisseling in ontwikkeling:
<https://github.com/minbzk/gdi-gegevensuitwisseling>

NORA visie op metadatamanagement:
https://www.noraonline.nl/wiki/Visie_metadatamanagement

NORA raamwerk gegevenskwaliteit:
https://www.noraonline.nl/wiki/Raamwerk_gegevenskwaliteit

Danny.Greefhorst@minbzk.nl

Werk sessie: Principles voor metadata voor gegevensuitwisseling





1.1. Gegevens die kunnen worden gedeeld zijn vindbaar, toegankelijk, interoperabel en herbruikbaar

Rationale	<p>Hergebruik van gegevens leidt tot meer transparantie, efficiëntie en innovatie. Voor het kunnen hergebruiken van gegevens zijn hun vindbaarheid, toegankelijkheid, interoperabiliteit en herbruikbaarheid zoals bedoeld met de FAIR-principes essentieel. Zij zorgen ervoor dat anderen in staat zijn om te begrijpen wat eigenschappen van de gegevens zijn, wat ze betekenen en of ze bruikbaar zijn voor hun eigen doeleinden. Ze zorgen er ook voor dat de gegevens eenvoudig verkregen kunnen worden.</p>
Implicaties	<ol style="list-style-type: none">1. Gegevens zijn voorzien van metagegevens, die zijn gepubliceerd in de stelselcatalogus, data.overheid.nl en als het geodata betreft ook in het Nationaal Geo Register.2. Gegevens en hun metagegevens zijn voorzien van wereldwijd unieke en stabiele identificaties (vindbaar).3. Metagegevens zijn rijk en bevatten ook informatie over de context, kwaliteit en karakteristieken van gegevens (vindbaar).4. Gegevens kunnen worden benaderd via standaard, open en algemeen implementeerbare protocollen (toegankelijk).5. Gegevens zijn beschreven in een gegevensmodel waarin hun syntax en betekenis zijn gedefinieerd en dat beschikbaar is een gestandaardiseerd open formaat zoals MIM (interoperabel).6. Metagegevens beschrijven gebruiksbependingen, -rechten en -plichten en de herkomst van de gegevens (herbruikbaar).



3.1. Gemeenschappelijke begripsvorming is het startpunt

Rationale	Taal is een belangrijke basis voor het creëren van wederzijds begrip. Het is tevens een essentiële basis voor gegevens, die een concrete invulling geven aan taal. Als we dus niet dezelfde begrippen hanteren, dan zullen de gegevens ook niet uitwisselbaar zijn. Afstemming over begrippen zou zo vroeg mogelijk moeten starten, nog voordat wet- en regelgeving wordt opgesteld. Dat wat in de wet- en regelgeving is opgenomen is nu eenmaal de formele basis voor gegevens en informatievoorziening en kan niet meer van afgeweken worden.
Implicaties	<ol style="list-style-type: none">1. In alle domeinen worden de daarin gehanteerde begrippen expliciet gemaakt.2. Begrippen worden gedeeld zodat deze vindbaar en toegankelijk zijn voor anderen.3. Begrippen worden geharmoniseerd en gestandaardiseerd waar mogelijk, zodat ze alleen van elkaar verschillen als er inherent andere betekenissen bij horen.4. Wet- en regelgeving wordt afgestemd op bestaande begrippen.5. In alle relevante documentatie, ontwerp en modellen in het domein wordt verwezen naar de begrippen, zodat ze beter vindbaar zijn en "by design" met elkaar samenhangen.



3.2. Metagegevens zijn begrijpelijk voor mensen

Rationale	Gebruikers willen op basis van metagegevens bepalen wat gegevens betekenen voor hun eigen gebruik en moeten deze daarvoor kunnen begrijpen. Specialisten hebben daarnaast ook een verantwoordelijkheid voor het onderhouden van metagegevens, en moeten deze metagegevens dus ook kunnen begrijpen om ze te kunnen aanpassen. Bepaalde metagegevens worden gegenereerd uit andere (hoger liggende) metagegevens en zijn daarmee automatisch op een hoger niveau gedocumenteerd.
Implicaties	<ol style="list-style-type: none">1. Metagegevens op alle niveaus zijn voorzien van een definitie of omschrijving en als dat niet voldoende is, ook van een aanvullende toelichting en/of voorbeelden.2. Metagegevens zijn geformuleerd in begrijpelijke taal, passend bij de doelgroep.3. Metagegevens zijn niet alleen beschikbaar in machineleesbare vorm, maar zijn voorzien van begrijpelijke documentatie.



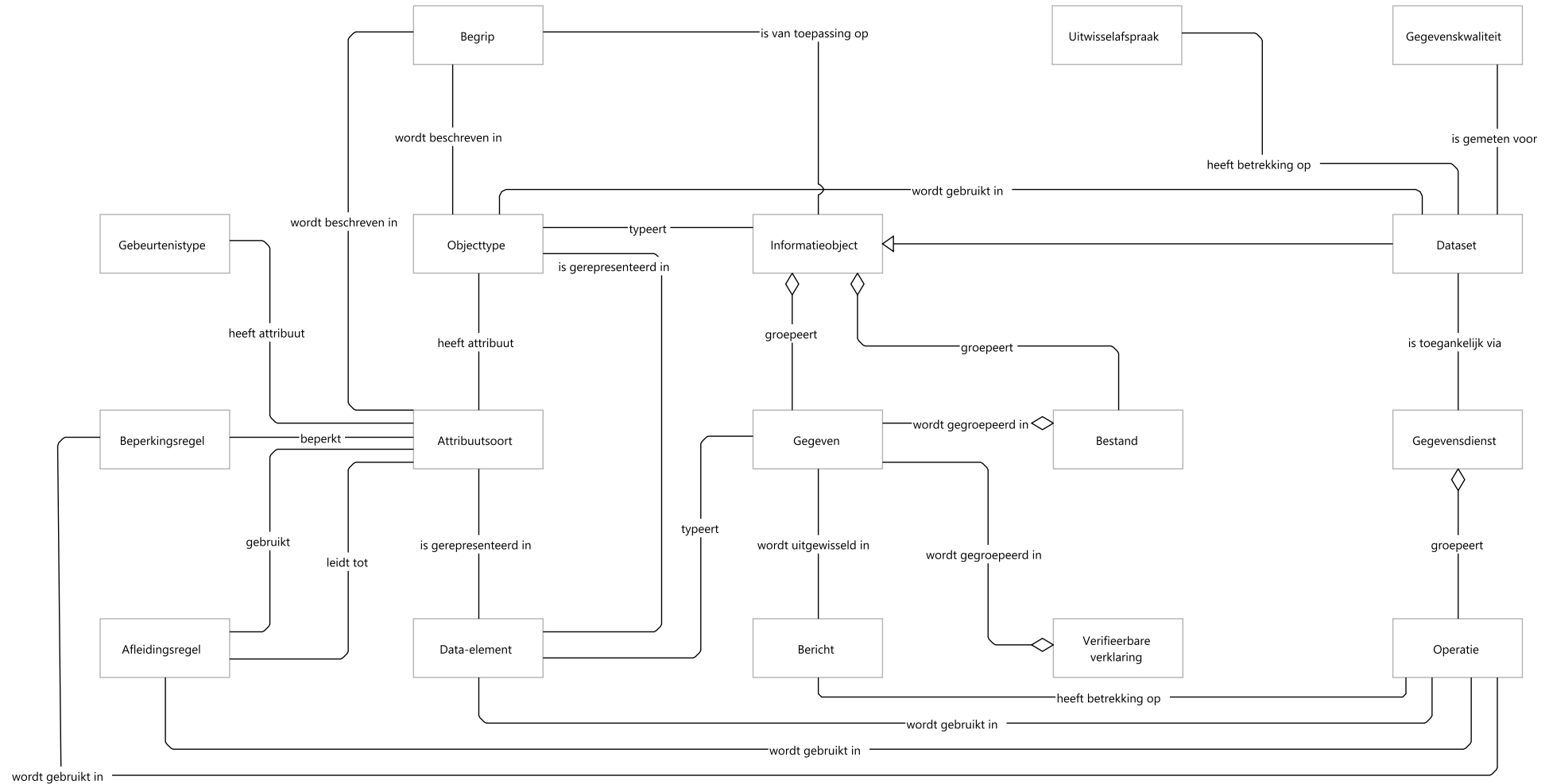
3.3. Gegevens worden contextrijk vastgelegd

Rationale	<p>Het is belangrijk om de context van gegevens te kennen om hun betekenis goed te begrijpen. Dat is in het algemeen het doel van metagegevens. Metagegevens worden echter vaak niet goed vastgelegd omdat gebruikers het nut ervan niet goed begrijpen of omdat het teveel inspanning kost om ze vast te leggen. Door context expliciet en geautomatiseerd vast te leggen in metagegevens wordt de afhankelijkheid van mensen voor het vastleggen van metagegevens beperkt. Het belang van expliciet maken van context in metagegevens is groter naarmate de afstand tot het initiële gebruik groter wordt.</p>
Implicaties	<ol style="list-style-type: none">1. Er wordt per gebruikcontext expliciet nagedacht welke metagegevens kritisch zijn.2. Metagegevens worden zoveel mogelijk gedurende een bedrijfsproces geautomatiseerd afgeleid uit de context en opgeslagen bij gegevens.3. Het vastleggen van metagegevens is geïntegreerd in de reguliere applicaties waar gebruikers mee werken.4. Metagegevens die niet automatisch kunnen worden afgeleid, worden expliciet door gebruikers aan de gecreëerde gegevens toegevoegd zoals trefwoorden, classificaties en omschrijvingen.



3.4. Metagegevens zijn aan elkaar verbonden

Rationale	Gegevens hebben metagegevens nodig om betekenis te krijgen en bruikbaar te zijn. Er zijn allerlei soorten metagegevens die verschillende aspecten van gegevens modelleren, zoals metagegevens die de betekenis beschrijven, metagegevens die de structuur beschrijven en metagegevens die beschrijven welke eenheden van gegevens beschikbaar zijn. Dit soort metagegevens moeten in relatie tot elkaar geïnterpreteerd worden, om echt waardevol te zijn.
Implicaties	<ol style="list-style-type: none">1. Modelelementen in informatie- en gegevensmodellen zoals objecttypes en attribuutsoorten refereren expliciet naar de gedefinieerde begrippen die ze representeren.2. In informatie- en gegevensmodellen wordt ook gerefereerd naar de beperkingsregels en afleidingsregels die van toepassing zijn.3. Bij metagegevens over datasets zijn ook metagegevens beschikbaar over de objecttypen en attribuutsoorten van de in de dataset aanwezige gegevens.4. Informatieobjecten zoals documenten en datasets zijn voorzien van begrippen zodat ze beter vindbaar zijn.5. Bij de uitgewisselde gegevens is een verwijzing aanwezig naar het gebruikte schema.6. Data-elementen in schema's (zoals JSON schema's en XML schema's) verwijzen expliciet naar de objecttypen en attribuutsoorten die ze representeren.7. Er is bij de documentatie van diensten of operaties in datadiensten beschreven welke beperkingsregels en afleidingsregels worden toegepast.



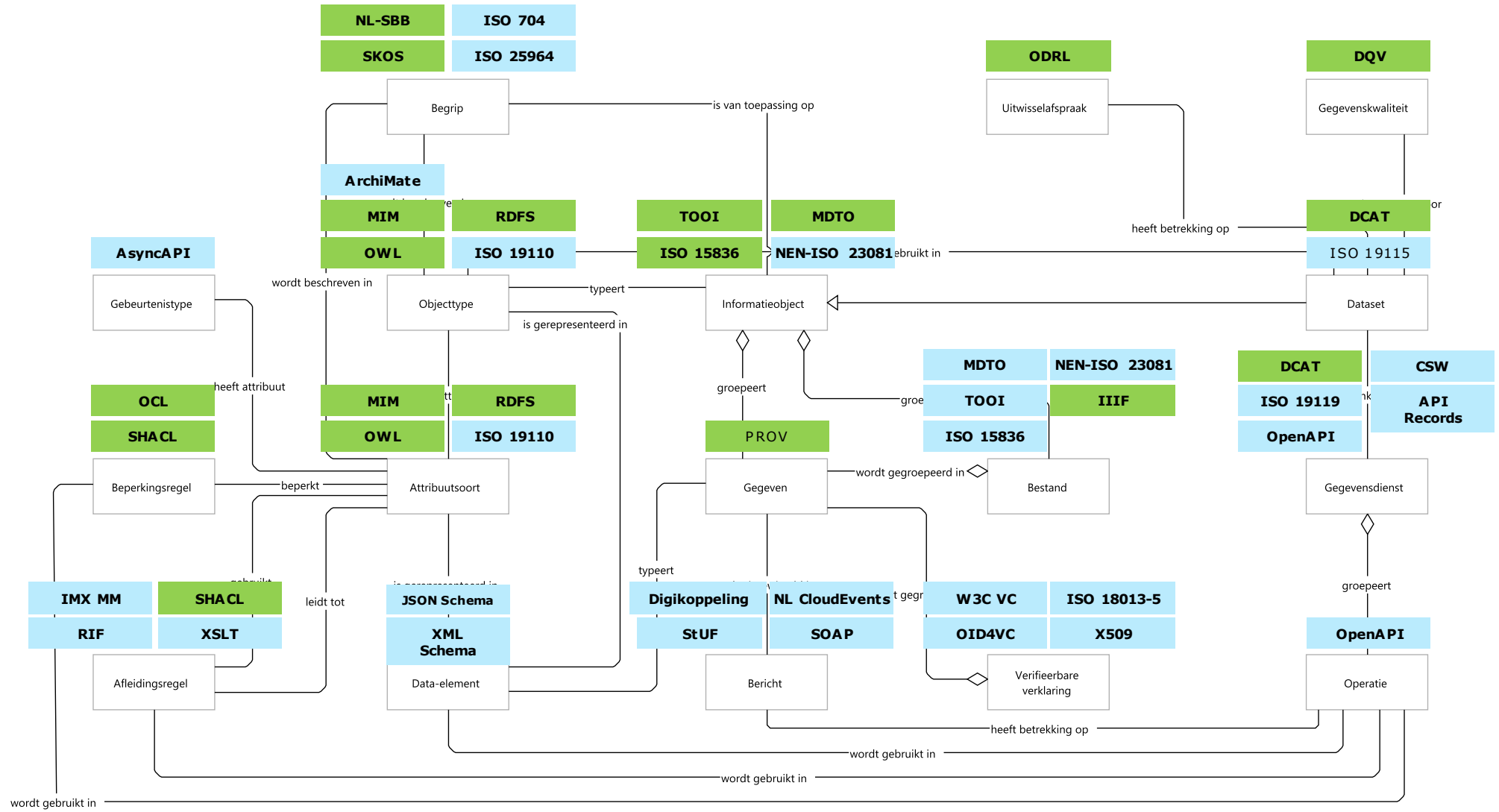


3.5. Metagegevens zijn beschikbaar als Linked Data

Rationale	Het is belangrijk dat metagegevens aan elkaar verbonden zijn om maximaal context te geven aan gegevens en ze echt waardevol te maken. Praktisch is het daarvoor nodig om verbindingen te leggen tussen verschillende vormen van metagegevens. Linked Data standaarden zijn specifiek gericht op het verbinden van gegevens en zijn geoptimaliseerd om gegevens te publiceren op het web. Er zijn allerlei Linked Data standaarden die ook specifiek gericht zijn op metagegevens.
Implicaties	<ol style="list-style-type: none">1. Metagegevens zijn beschikbaar in de wereldwijd veelgebruikte Linked Data vocabulaires voor metagegevens zoals SKOS, Dublin Core, RDFS, OWL, SHACL, DCAT, PROV-O, DQV, ODRL en IIIF.2. Nederlandse standaarden voor metagegevens bieden (ook) een Linked Data vocabulaire, zoals dat ook is ontwikkeld voor bijvoorbeeld NL-SBB, MDTO, TOOI en MIM.3. Voor wereldwijde standaarden voor metagegevens waarvoor geen Linked Data vocabulaires aanwezig zijn wordt een standaard Nederlandse vocabulaire ontwikkeld, zoals voor de OpenAPI Specification en ArchiMate.4. Op de raakvlakken tussen de standaarden wordt middels webadressen (IRI's) verwezen naar gerelateerde metadatastandaarden.5. Metagegevens worden gepubliceerd in landelijke catalogi zoals data.overheid.nl, developer.overheid.nl en de stelselcatalogus, die deze (ook) als Linked Data ontsluiten.



**Linked Data
standaard**





3.6. De kwaliteit van gegevens is beschreven conform het landelijke raamwerk gegevenskwaliteit

Rationale	Het is belangrijk dat gebruikers eenvoudig inzicht kunnen krijgen in de kwaliteit van gegevens, zodat zij begrijpen of deze aansluiten bij hun eigen gebruik. Door de naamgeving en inhoud van dimensies voor gegevenskwaliteit te standaardiseren, zijn deze beter begrijpelijk voor afnemers. Het NORA raamwerk gegevenskwaliteit is gebaseerd op wereldwijde ISO standaarden, breed afgestemd binnen de Nederlandse overheid en is ook de basis voor het Federatief Datastelsel.
Implicaties	<ol style="list-style-type: none">1. In de metagegevens van datasets zijn beschrijvingen van de normen en gerealiseerde kwaliteit beschreven conform de structuur van het NORA raamwerk gegevenskwaliteit.2. Er is liefst ook een machineleesbare versie van de kwaliteit van de gegevens beschikbaar conform de DQV standaard.



NORA raamwerk gegevenskwaliteit

Juistheid	Thematische juistheid	Classificatie juistheid	Positionele juistheid	Kwantitatieve juistheid		
Compleetheid	Dataset compleetheid	Object compleetheid	Over compleetheid	Historie compleetheid	Metadata compleetheid	Ruimtelijke dekking
Validiteit	Formaat validiteit	Domein validiteit	Geometrische validiteit			
Consistentie	Logische consistentie	Referentiële integriteit	Identificeerbaarheid	Homogeniteit	Topologische consistentie	
Actualiteit	Update frequentie	Versheid				
Precisie	Opslagprecisie	Geometrische precisie	Statistische precisie			
Plausibiliteit	Authenticiteit	Reputatie	Bewijsbaarheid	Representativiteit		
Traceerbaarheid	Herleidbaarheid	Reproduceerbaarheid	Gebruiks inzicht			
Begrijpelijkheid	Leesbaarheid					



5.1. Informatieproducten zijn herleidbaar naar de onderliggende gegevens en regels

Rationale	Het is belangrijk dat afnemers van informatieproducten kunnen begrijpen waar de gegevens die zij afnemen op zijn gebaseerd. Hierdoor kunnen ze bepalen hoe de resultaten tot stand zijn gekomen, of het resultaat aansluit bij hun informatiebehoefte en of de onderliggende gegevens voldoende betrouwbaar zijn. Het geeft ze vertrouwen in de mate waarin het informatieproduct aansluit bij hun eigen gebruik. De kwaliteit van de onderliggende gegevens bepalen namelijk vooral de kwaliteit van het informatieproduct dat erop is gebaseerd.
Implicaties	<ol style="list-style-type: none">1. De afleidingsregels zoals gebruikt voor het samenstellen van het informatieproduct zijn toegankelijk voor de afnemer.2. In metagegevens van informatieproducten is expliciet gemaakt welke administraties zijn gebruikt.3. Liefst zijn ook de specifieke onderliggende gegevens die zijn gebruikt bij de afleiding toegankelijk voor de afnemer.4. Afleidingsregels zijn onder versiebeheer zodat ook historisch inzicht kan worden gegeven in hun werking.

Schuif eens aan bij de NORA Familie

meedenken - kennis opdoen - oplossingen delen

Informatie:

noraonline.nl

[lijst architecturen \(noraonline\)](#)

[Sessies op DADD](#)

Contact:

nora@ictu.nl

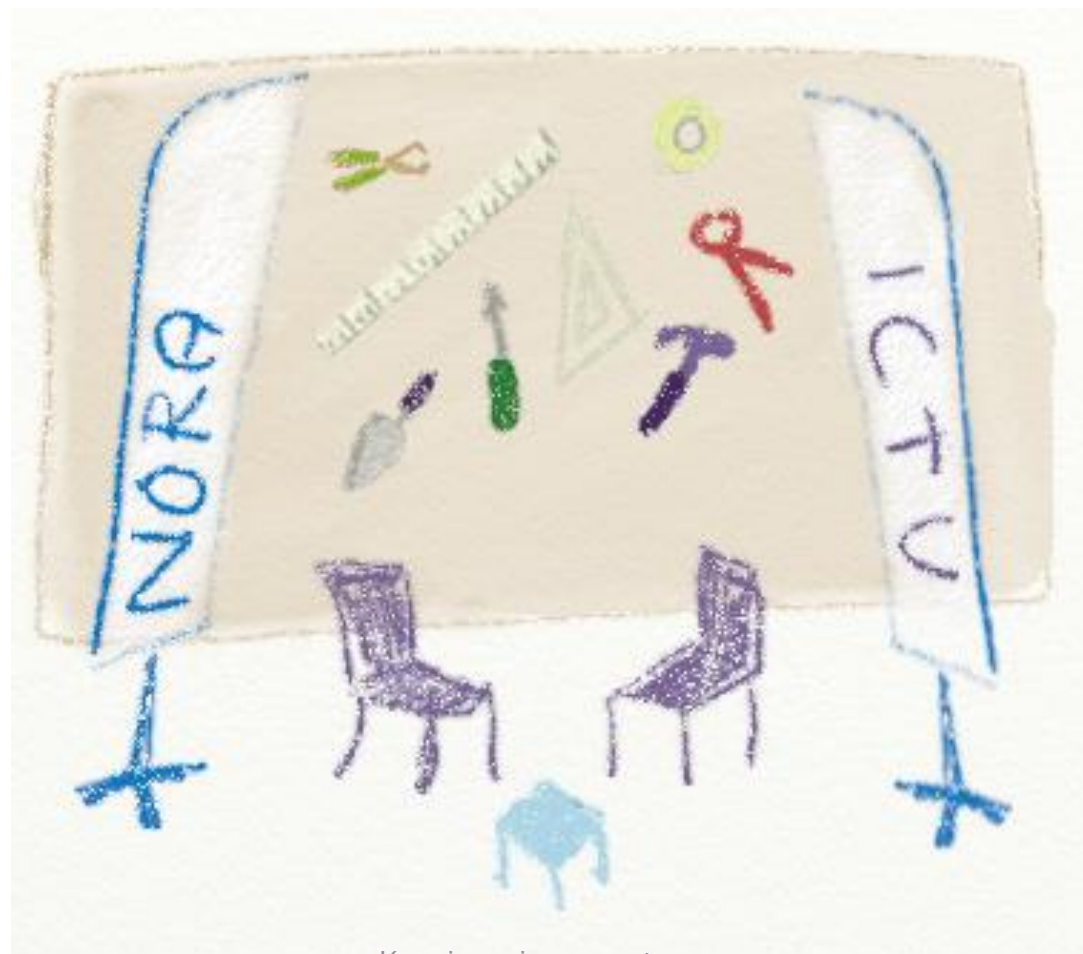
[NORA Beheer](#)

[@NORArchitectuur](#)

Deelname:

[NORA Gebruikersraad](#)

[Open Huis van de Architectuur](#)



Blijf op de hoogte:

[Nieuws & Agenda](#)

[Persoonlijke volglijst](#)

[NORA Familienieuws](#)

[RSS-feeds NORA](#)

Social media:

[X \(uitfaseren\)](#)

[LinkedIn](#)

[Mastodon](#)

Kennissessie 5 maart 2024