

MIM Community

Semantische Analyse

Hein Corstens

6 juli 2023

Inhoud

1. Probleemstelling
2. Semantische analyse in T3D (Totaal driedimensionaal)
3. Hoe verder: harmoniseren van definities en ontology alignment
4. Terugkoppeling naar MIM
5. Afsluiting



0. Als aftrap: een drietal stellingen

1. Een semantische basis is een noodzakelijke voorwaarde voor een goede informatievoorziening.
2. Aandacht voor semantiek leidt tot inzicht voor beheerders en gebruikers in waar ze mee bezig zijn en daarmee tot een betere informatievoorziening.
3. De oplossing voor gegevensincongruentie ligt niet in een brute gelijkschakeling van definities, en ook niet in het oppervlakkig mappen op basis van verwante begrippen, maar in een zorgvuldig en langdurig proces van het afstemmen van ontologieën ('ontology alignment').

1. Probleemstelling

Als het gaat om gegevensmodellering is er vaak sprake van:

- verschillende termen voor dezelfde objecttypen (synoniemen)
- verschillende definities voor dezelfde objecttypen
- gelijke termen voor verschillende objecttypen (homoniemen)
- gelijke definities voor verschillende objecttypen
- definities en termen die ONGEVEER gelijk zijn
- definities en termen die TEGENSTRIJDIG zijn
- enzovoorts

Laten we wat voorbeelden bekijken

Bank



Bank

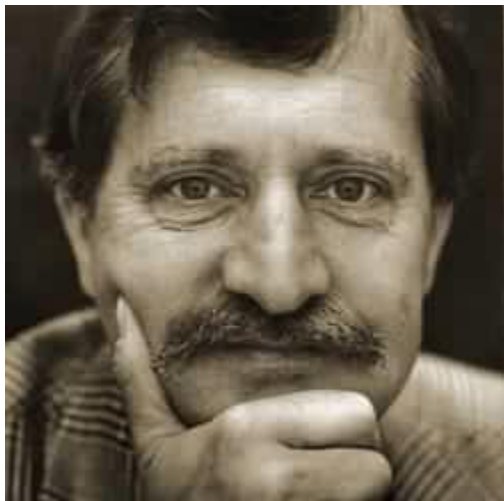
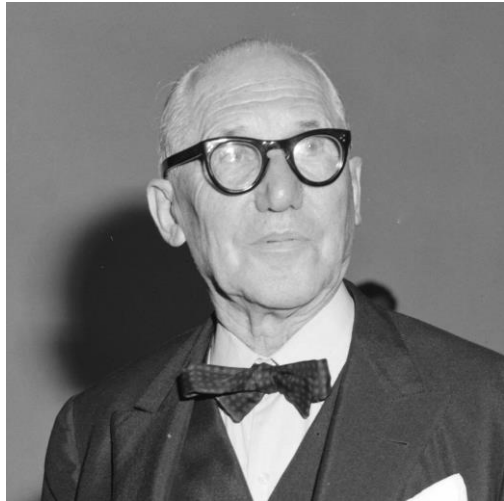
- databank
- vacaturebank
- zandbank
- toonbank
- draaibank
- mistbank

enzovoorts

luchtvaartobservatieligbank



architect



6-7-2023

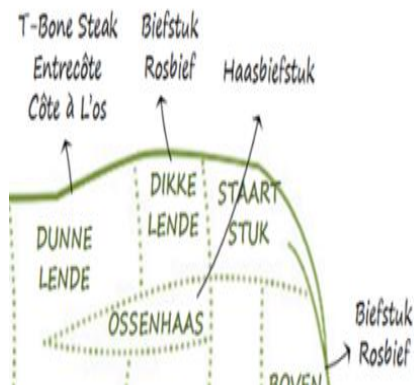
rosbief

- geroosterd mager rundvlees
- NIET: paardenbief



en:

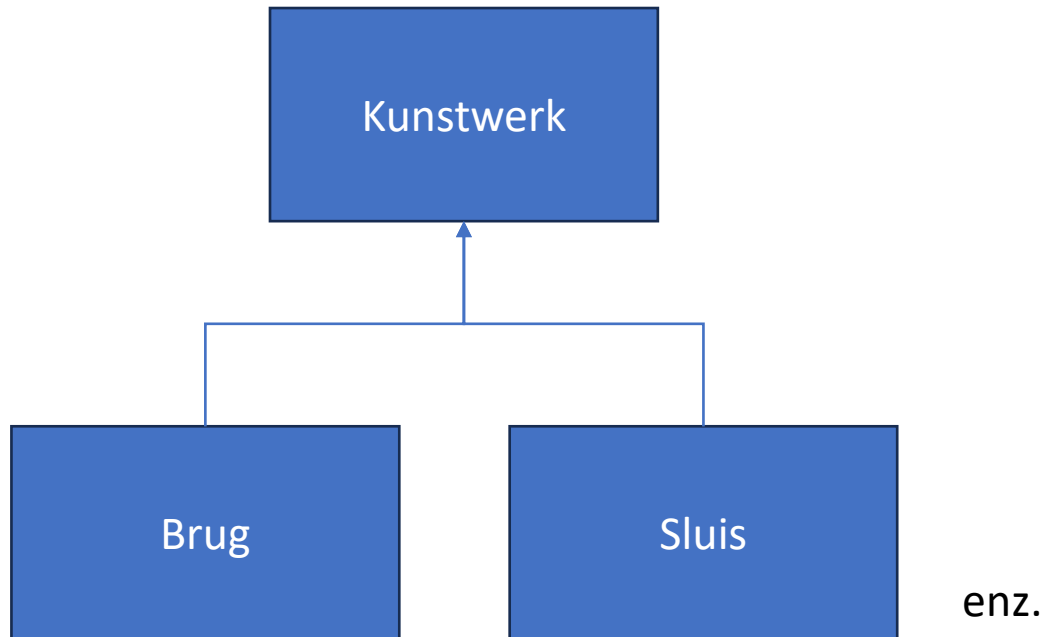
- biefstuk van de haas is géén biefstuk van de haas



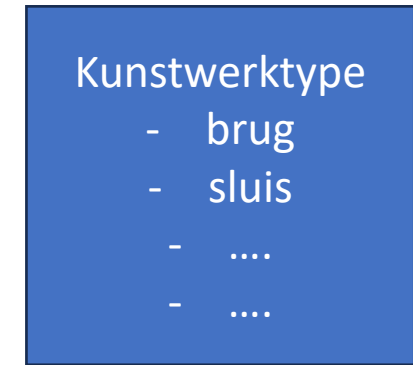
sexe/sekse

- tot 1995: sexe of sekse
- na 1995: alleen sekse
- M/V → N/M/V/I

kunstwerk



Subtypering



Enumeratie



Bord of kom?



Camper: woning én auto



Het miezert: het regent én het regent niet



ergernis en synoniemen



ergernis enz. (van Dale)

- ergernis: de toestand van geërgerd te zijn; < - - tot ontstemming geprikkeld zijn; (ontstemd: enigszins boos, niet te spreken)
- boosheid: (boos: ongunstig gestemd over iets)
- woede: grote kwaadheid
- verontwaardiging: ergernis
- wrevel: bitterheid, bedwongen spijt
- irritatie: geprikkelde stemming; = ergernis
- frustratie: gevoel van teleurstelling
- teleurstelling: het teleurgesteld-zijn (het niet beantwoord zijn aan een verwachting)

Over ergernis gesproken:

Nathan Ducastel, directeur VNG Realisatie (iBestuur 47, juni 23):

Kraan dicht en dweilen

Ducastel: "Wat mij tijdens een recent studiebezoek aan de KSZ duidelijk werd is dit: we moeten onze huidige informatie-kundige begrippen en hun componenten opnieuw definiëren met een unieke betekenis. Dan ontstaat een informatie-kundig landschap waarin iedereen dezelfde taal spreekt. Dit bevordert de interoperabiliteit van gegevens."

Hier geloven wij
dus niet in!

"We moeten de kraan draaien en gaan dweilen. Wet voor wet opschonen vanuit één informatiekundige visie en één informatiemodel. Te beginnen met de begrippen die nu de meeste verwarring veroorzaken in de uitvoering. Vervolgens die definities vastleggen en juristen, bestuurders en politici een eed laten zweren dat ze nooit meer aan nieuwe definities beginnen. Ook niet wanneer de politiek probeert hen te dwingen."

Samenvattend: verschillen en overeenkomsten in:

1. Terminologie
2. Syntax
3. Semantiek
4. Semiotiek (pragmatiek)

2. Semantische analyse (T3D-project)

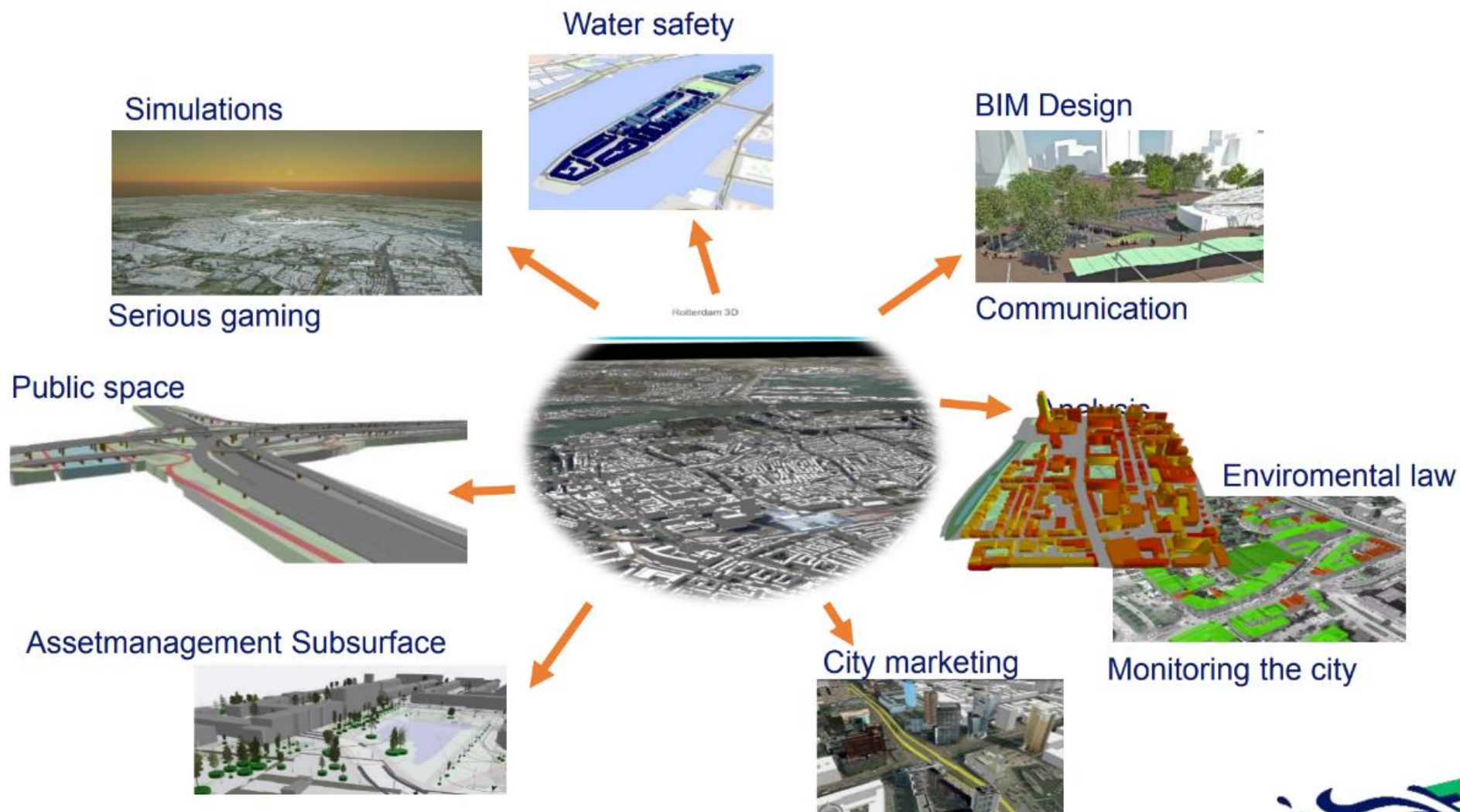
2.1 Het T3D-programma

2.2 Semantische analyse in T3D

2.1 T3D

- Programma van VNG+gemeenten Amsterdam, Rotterdam en Den Haag (2021-2023)
- Doel: Verbetering van de 3D informatievoorziening door gemeenten
- Middelen: usecases, proefopstelling, beproevingen
- Resultaat: ruim 30 producten en rapportages. Zie <https://alkem.io/t3d> en [eindrapport](#).

3D Stadsmodel: gebruiksmogelijkheden



Fysieke stad

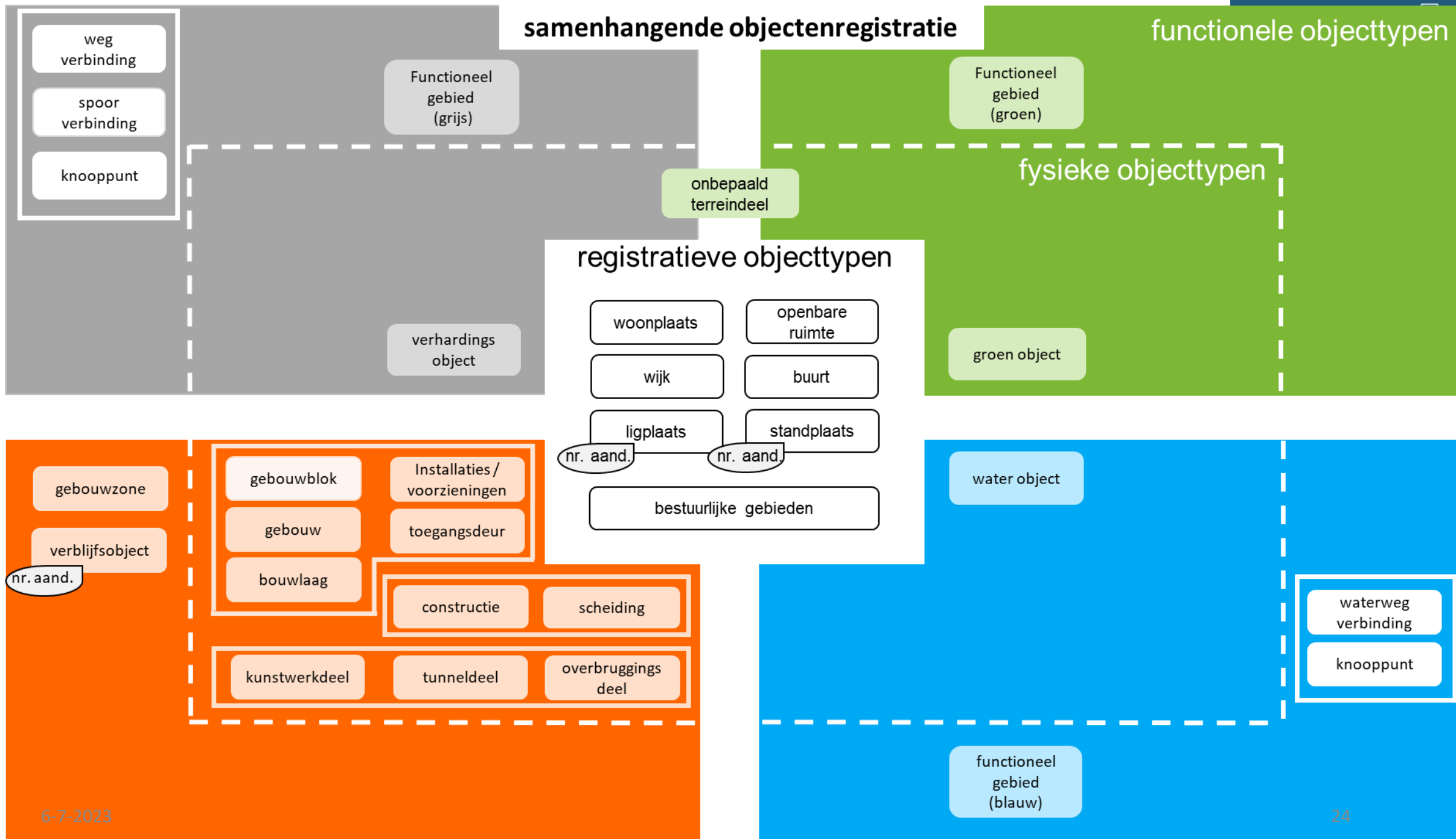


Digitale stad

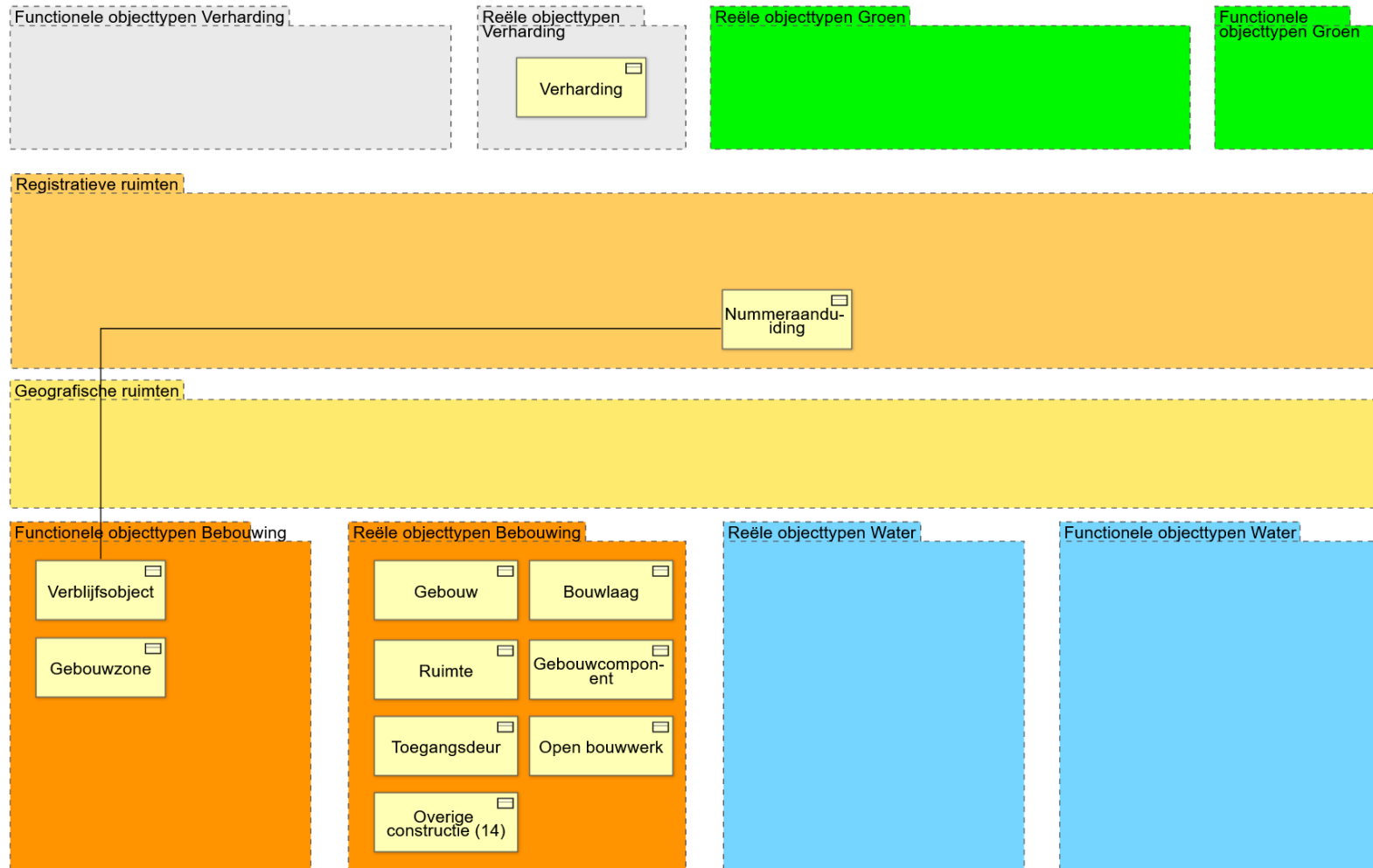


Nodig: integraal gegevensmodel (of niet?)

- Begrippenmodel
- Conceptueel model
- Logisch(e) model(len)
- Technisch(e) model(len)

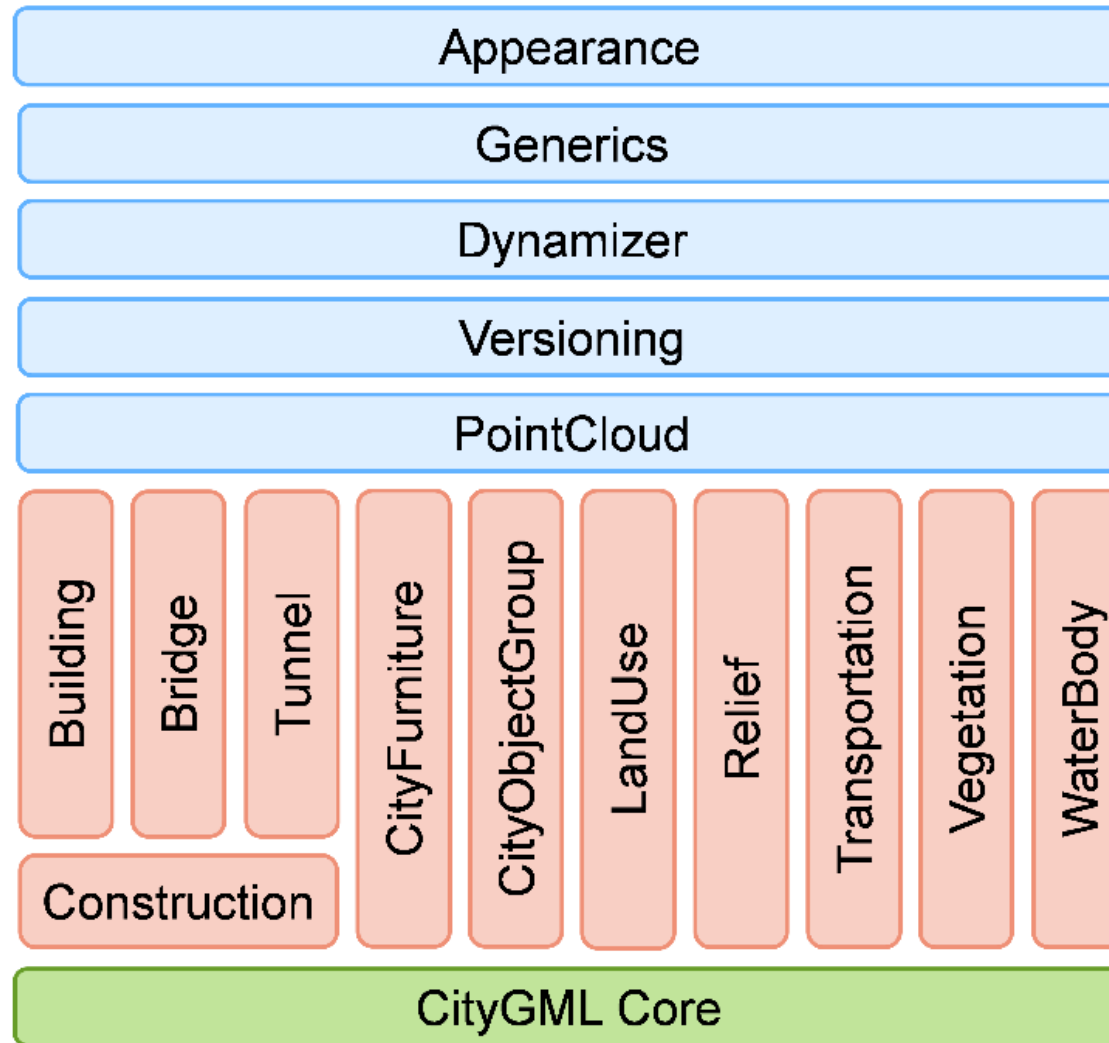


Begrippenmodel T3D

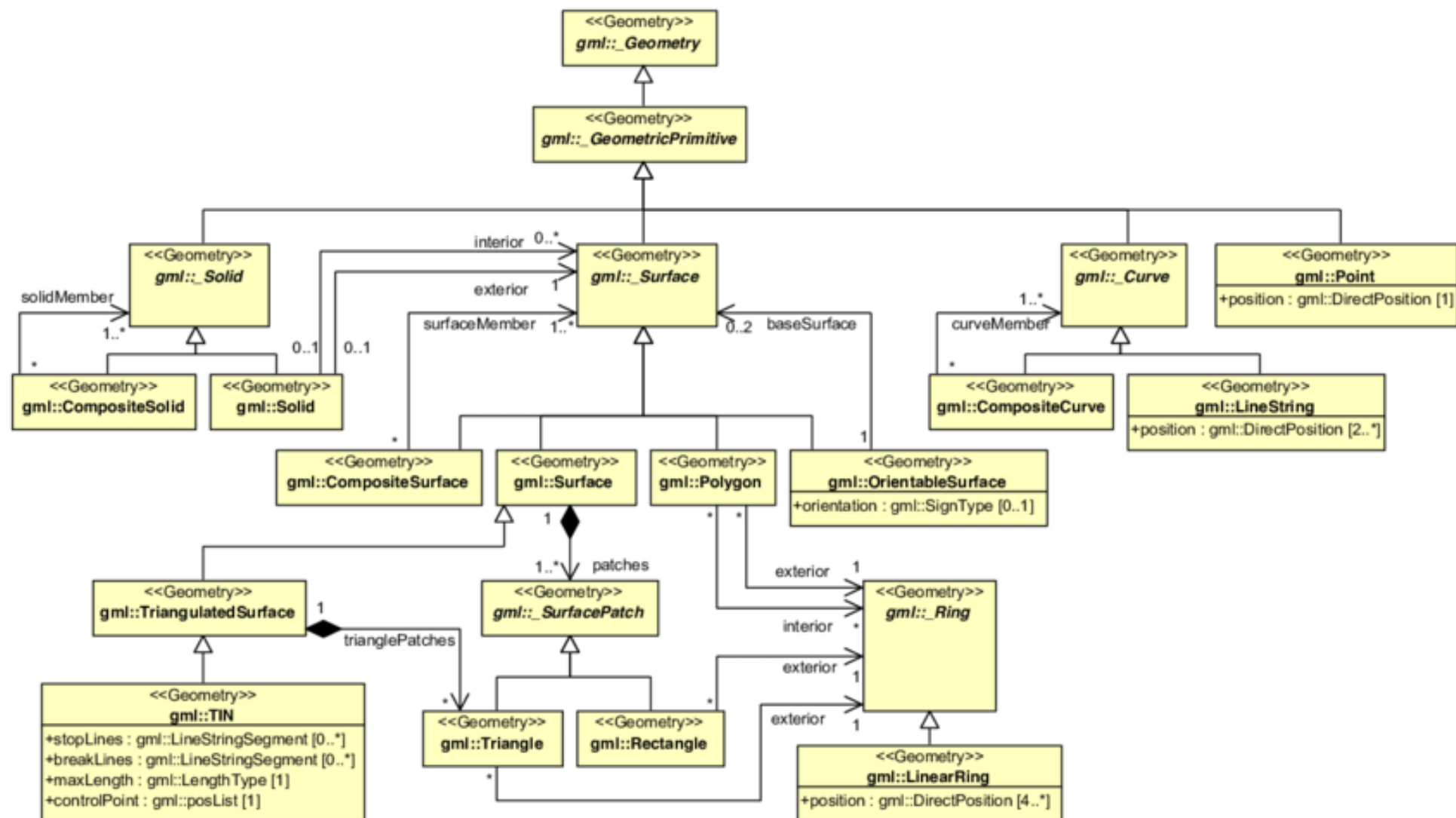


ArchiMedes 08-12-2022 10:55:04 CET

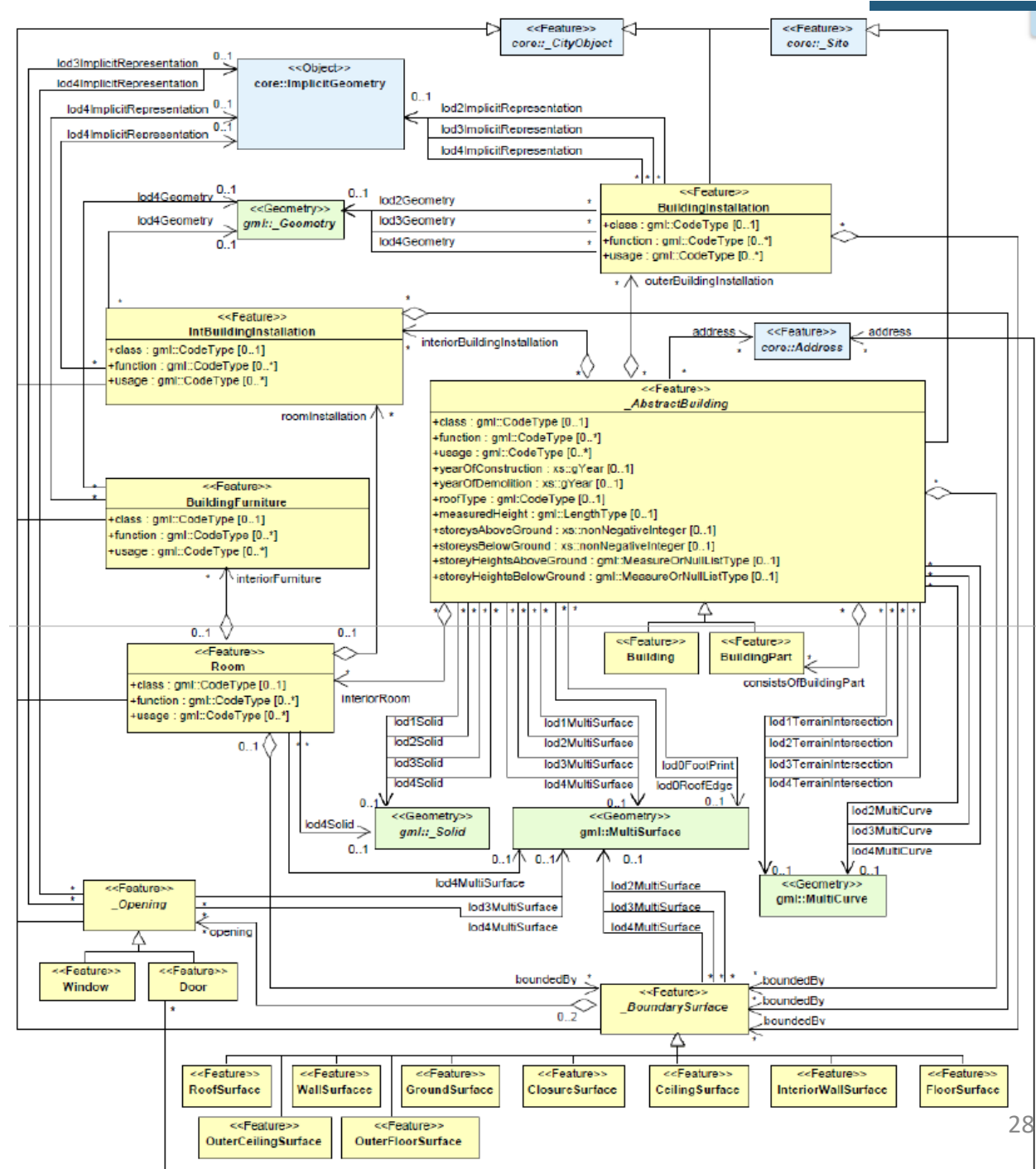
Basis: CityGML



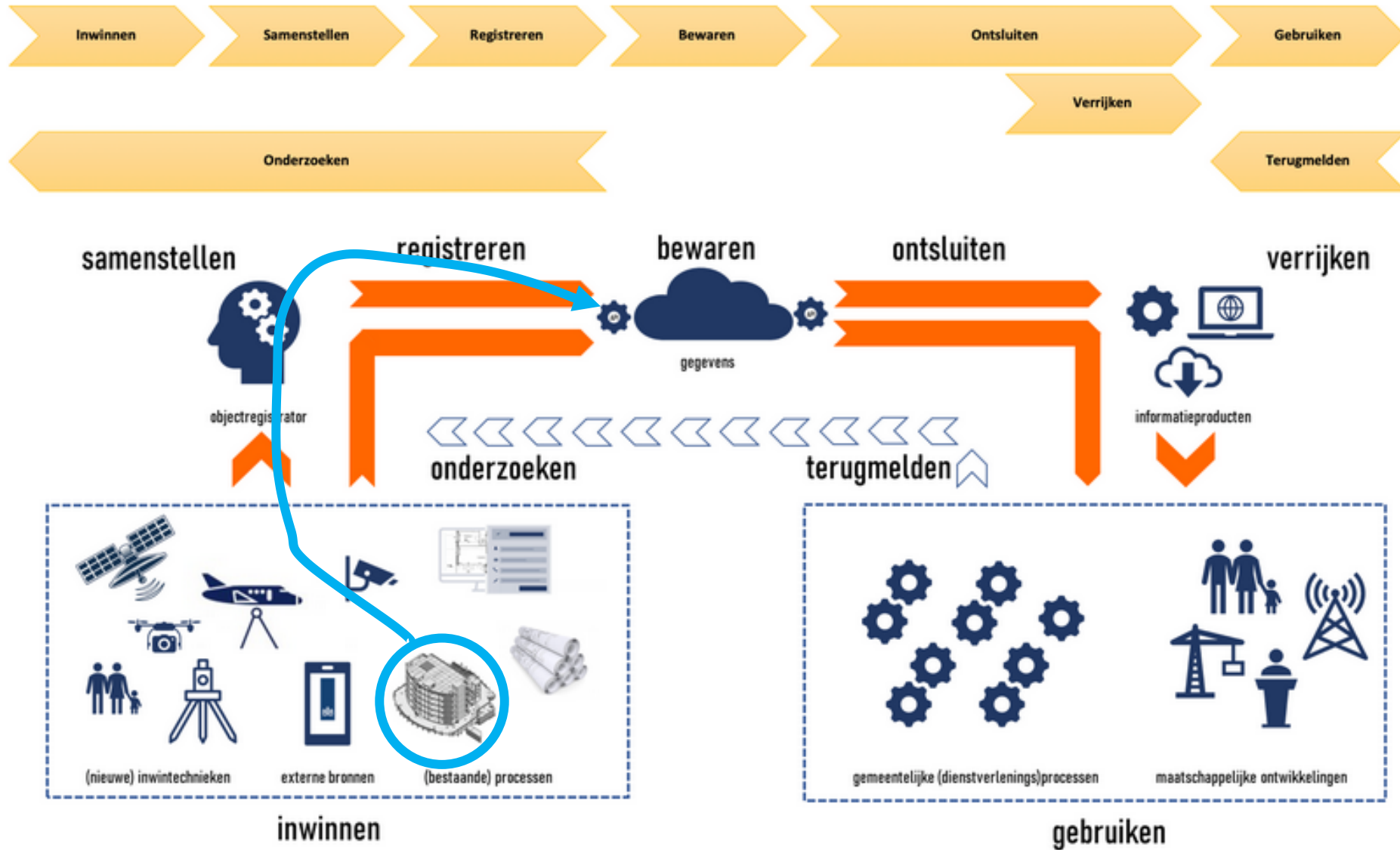
CityGML-geometry model



CityGML: Building model

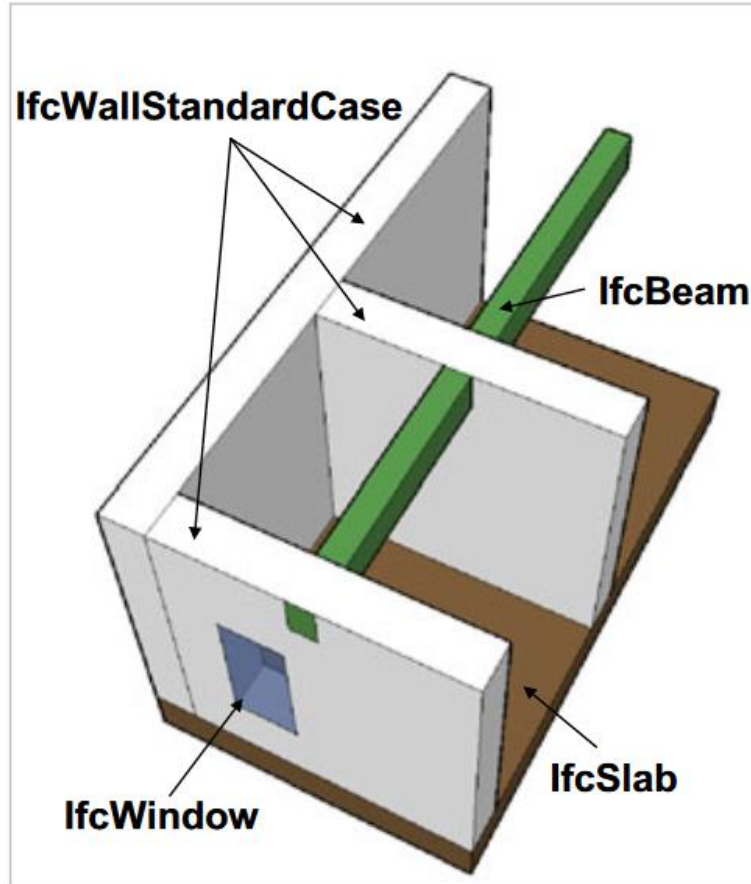


T3D-keten



Building Information Modeling (e.g. IFC)

e.g. Constructive Solid Geometry



Semantic 3D City Modeling (e.g. CityGML)

Boundary Representation

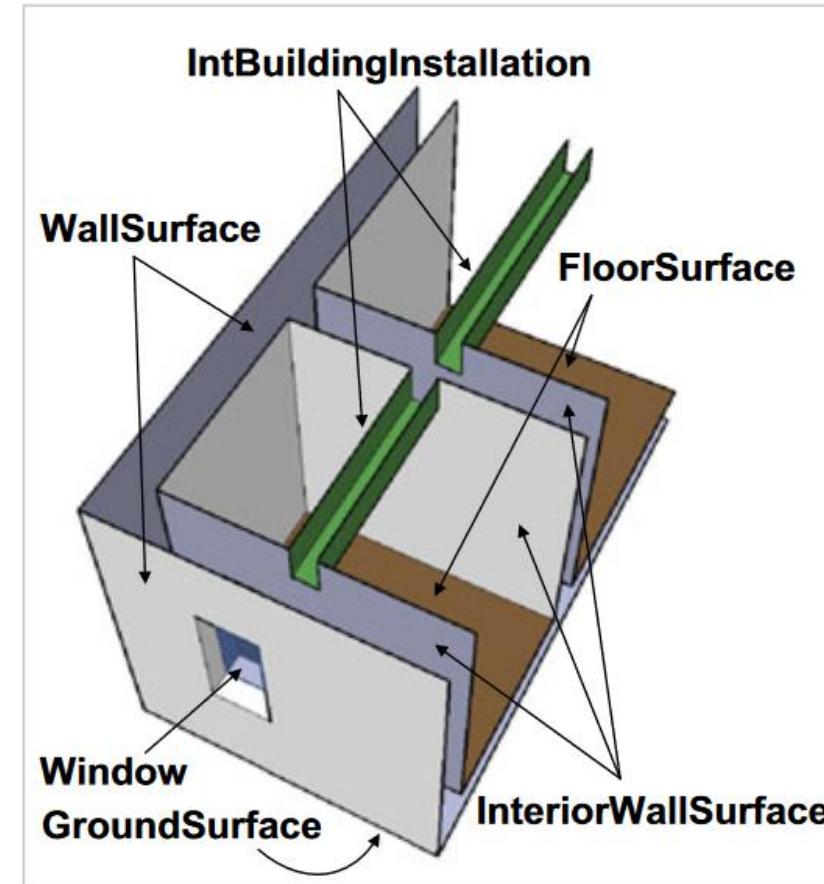
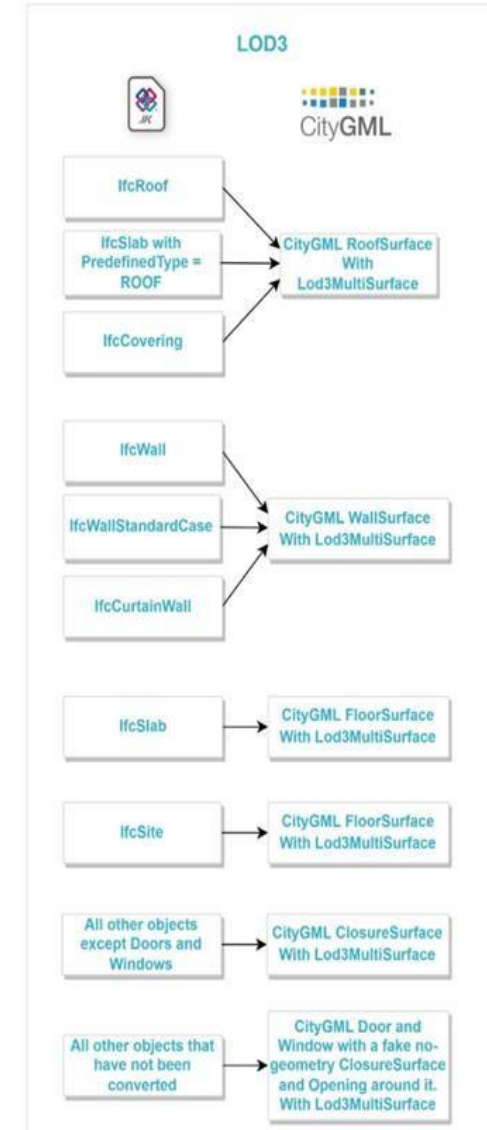
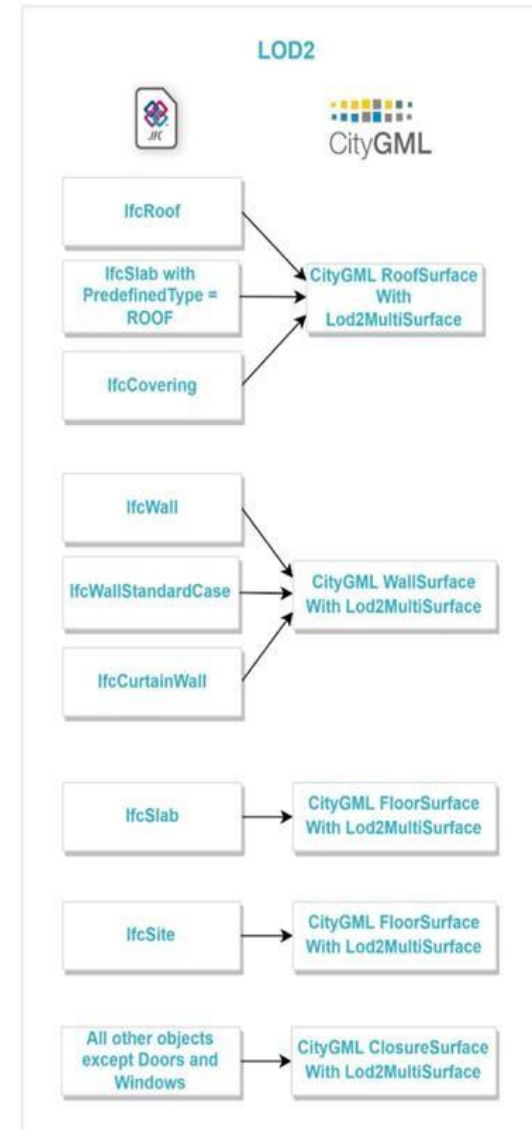
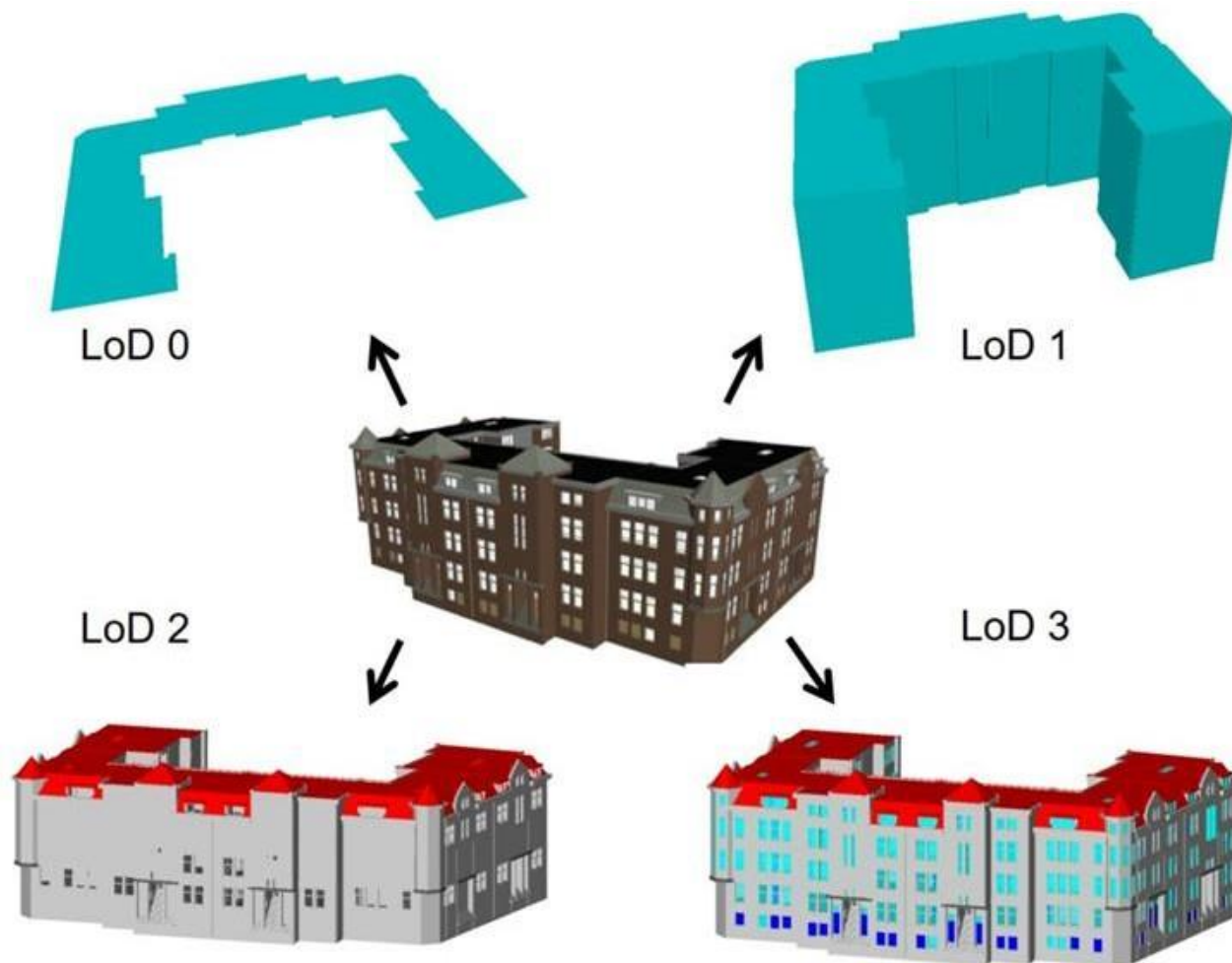


Fig. 34.7 Geometry modeling paradigms predominantly applied in BIM and semantic 3D city modeling (Nagel et al. 2009)



2.2 Semantische analyse T3D

- Doel: inzicht in overeenkomsten en verschillen in terminologie en definitie van relevante begrippen in T3D
- Achterliggend doel: afstemming tussen begrippen in verschillende domeinen
- Aanpak
 - Inrichting tool (i.c. de GWSW-server)
 - Inventarisatie en analyse concepten (Gebouw, Weg, Kolk, Boom, Geometrie)
 - Conclusies en aanbevelingen

GWSW

Soorten Samenstelling Collecties

Zoek...

Soorten

Vouw open

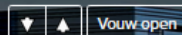
- Apparatuur
- Bouwwerk
- Constructieonderdeel
- Kabel
- Leiding
- Materie
- Put
 - Aansluitput
 - Beerput
 - Blinde put
 - Brandput
 - Drainageput
 - Filterput
 - Infiltratieput
 - Kolk
 - Loze put
 - Ontlastput
 - Opvangput
 - Rioolput
 - Bijzondere putconstructie
 - Valput
 - Doorspoelput
 - Doorspuitput
 - Inspectieput
 - Kruisingsput
 - Lozingsput
 - Mantogankelijke put
 - Overstortput
 - Pompput

Pompput

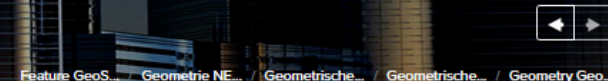
Beschrijving

Naam	Pompput
URI	Pompput
Definitie	Een rioolput met een inrichting bestemd voor het verpompen van afvalwater.
Zie ook	https://data.gsw.nl/img/pompput2.jpg [NEN 3300:1996] rioolput met een inrichting voor het verpompen van afvalwater
Datum start/wijz	2006-02-20 / 2007-11-12
Heeft representatie	NLCS_GWA_RIOOLPUT_POMPPUT (1719) [max=1] NLCS_HWA_RIOOLPUT_POMPPUT (1848) [max=1] NLCS_ITR_RIOOLPUT_POMPPUT (2005) [max=1] NLCS_PERSLEIDING_RIOOLPUT_POMPPUT (2076) [max=1] NLCS_SWA_RIOOLPUT_POMPPUT (2126) [max=1] NLCS_VACUUMLEIDING_RIOOLPUT_POMPPUT (2216) [max=1] NLCS_VWA_RIOOLPUT_POMPPUT (2269) [max=1] Sym_Pompput [max=1]
Functie	Verpompen van afvalwater +Leidingen aansluiten +Rioolleidingen aansluiten
Uitvoering	+Vertikale waterdichte constructie +Gegraven of geboorde kokervormige diepte
Heeft kenmerk	+Aantal woningen [max=1] +Aantal i.e. bedrijven [max=1] +Aantal i.e. recreatie [max=1] +Adresgegevens [max=1] +Afvoerend oppervlak [max=12] +Begindatum [max=1]

Soorten



IfcObjectPlacement IFC



- Ding (datamodel T3D)
 - ▷ ○ Activiteit
 - ▲ ○ Fysiek object
 - ▷ ○ Bouwwerk
 - ▷ ○ Deelconstructie
 - GeoObject NEN 3610
 - ▷ ○ Levensvorm
 - ▷ ○ Open water object
 - ▷ ○ Put
 - ▷ ○ Wasteterminal
 - ▷ ○ Informatiedrager
 - ▷ ○ Kenmerk
 - ▲ ○ Ruimte
 - Feature GeoSPARQL
 - ▷ ○ GM_Object
 - Geometrie NEN 3610:2022
 - GeometrischeEntiteit NEN2660-1
 - GeometrischeEntiteit NEN2660-2
 - Geometry GeoSPARQL
 - ▲ ○ IfcObjectPlacement IFC
 - IfcGridPlacement
 - IfcLinearPlacement
 - IfcLocalPlacement
 - IfcProductDefinitionShape IFC
 - IfcShapeRepresentation IFC
 - IfcTopologyRepresentation IFC
 - ImplicitGeometry CityGML
 - Metric space OntoMathPro
 - SpaceBoundary CityGML
 - ▷ ○ Space CityGML
 - Verblijfsobject
 - ▷ ○ Topologisch element
 - ▷ ○ Element NEN 2660

Beschrijving

Naam	IfcObjectPlacement IFC
URI	IfcObjectPlacementIFC
Definitie	IfcObjectPlacement is an abstract supertype for the special types defining the object coordinate system. The IfcObjectPlacement has to be provided for each product that has a shape representation. The object placement can be given: absolute: by an axis2 placement, relative to the world coordinate system, relative: by an axis2 placement, relative to the object placement of another product, by grid reference: by the virtual intersection and reference direction given by two axes of a design grid, linear placement: by distance along a curve, with possible offsets. Definities vanuit supertypes: NONE
Datum start/wijz	2023-02-03 / 2023-02-03
Functie	+Vastleggen locatie
Is kenmerk van	Building IFC [inverse]
	Vegetation IfcGeographicElement [inverse]
Heeft supertype	Ruimte
	Geometrie
	Geometrie, locatie
Heeft subtype	IfcGridPlacement
	IfcLinearPlacement
	IfcLocalPlacement
Model details	
Heeft scope	Cof_GEO
Auteur start	Projectteam T3D Ontologie

Gebouw

6-7-2023

Gebouw/IFC-CityGML

<i>Standaard</i>	<i>Kenmerk</i>	<i>CityGML</i>	<i>IFC</i>
Naam		Building	Building
Functie	Faciliteren fabricage of diensten		
	Beschermen		
	Onderbrengen		
	Onderbrengen mensen		
	Onderbrengen dieren		
	Onderbrengen goederen		
Structuur	Duurzaam verbonden		
	Artefact		
	Reëel		
	Discreet		
	Direct verbonden		
	Direct/indirect verbonden		
	Samengesteld		
Uitvoering	Geheel/gedeeltelijk omsloten met wanden		
	Geheel omsloten met wanden		
	Menstoegankelijk		
	Vrijstaand		
	Constructief zelfstandig		
	Overdekt		
	Gefundeerd		
	Heeft objectplaatsing		
	Heeft pilaren		
	Bovengronds		
	Ondergronds		
	Afsluitbaar		

Boom



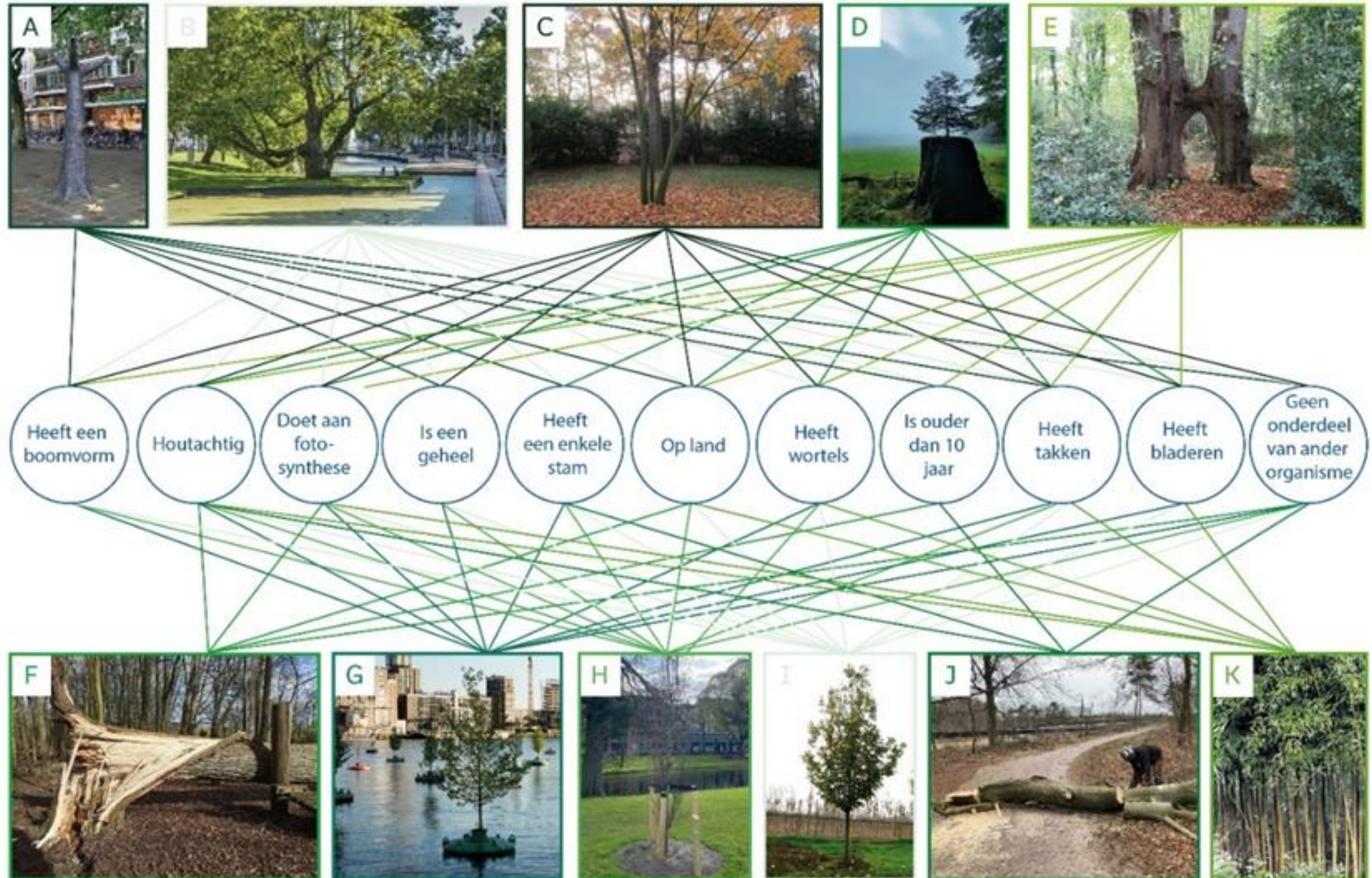
meerstammige boom aan het Weena in Rotterdam

- in Rotterdam: Boom
- géén IMBOR:Boom
- wél een IMGEO:Boom
- betere aansluiting bij:
IMBOR:Struik, maar er is
geen IMGEO:Struik

Boom

<i>Standaard</i>	<i>Kenmerk</i>	<i>IMGEO</i>	<i>IMBOR</i>
Naam		Boom	Boom
Structuur	Reëel		
	Discreet		
	Duurzaam verbonden		
	<u>Aangepland</u>		
	Anatomisch		
	Organisme		
	Meercellig		
	Natuurlijk ontstaan		
	Topografisch		
Uitvoering	Solitair		
	Geen onderdeel van andere beplanting		
	Bedekking		
	Markant		
	Houtachtig		
	Vertakkend		
	Op land		
	Voortplantingsorganen		
	Relatief hoog		
	Beperkte omvang		
Technology	Fotosynthese		
Functie	Voedselbron		
	Beschermend		
	Nestgelegenheid biedend		
	Grondstof		
Samenstelling	Heeft deel		Wortelgestel (min 1 max 1) Stevige houtachtige stam (min 1 max 1)

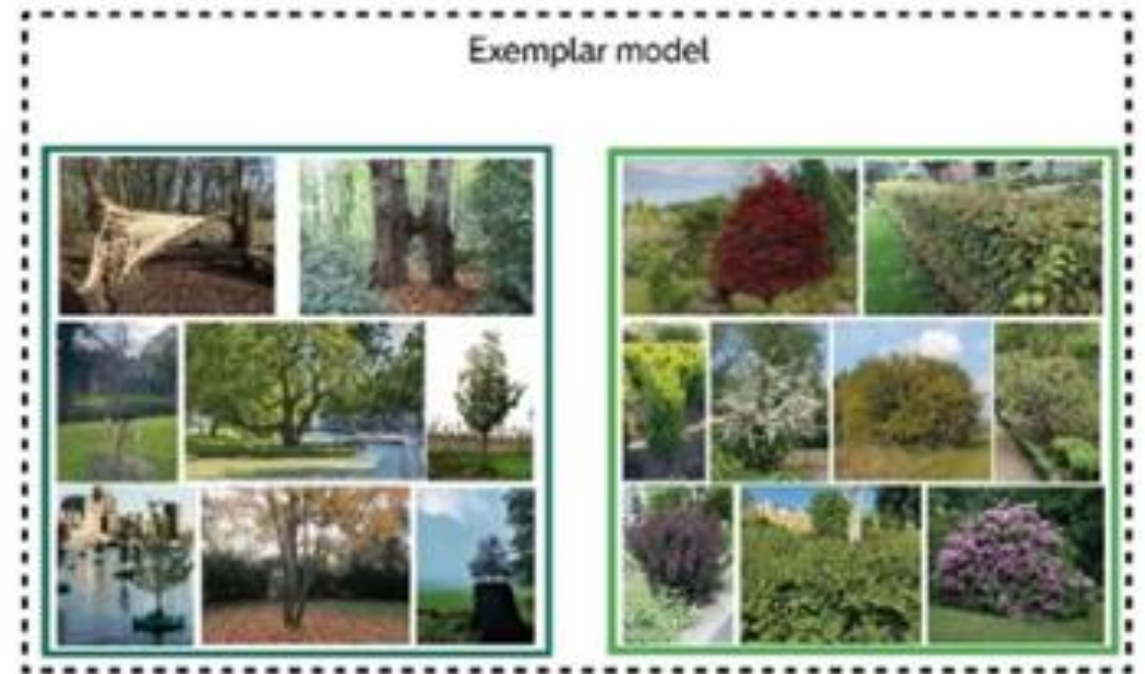
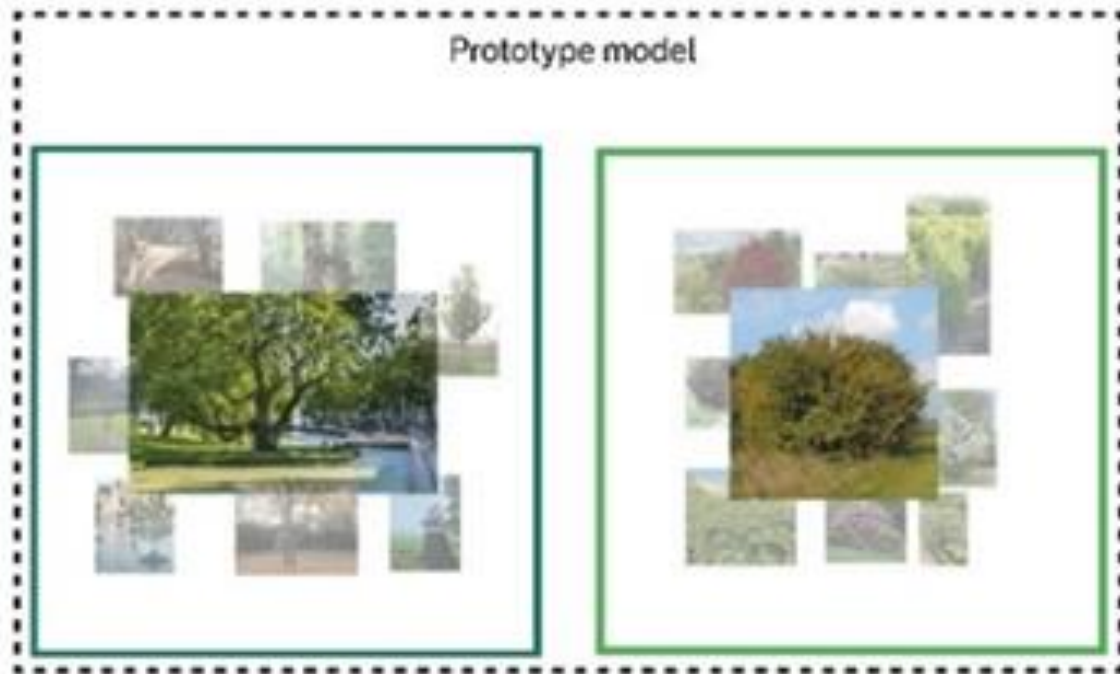
Boom



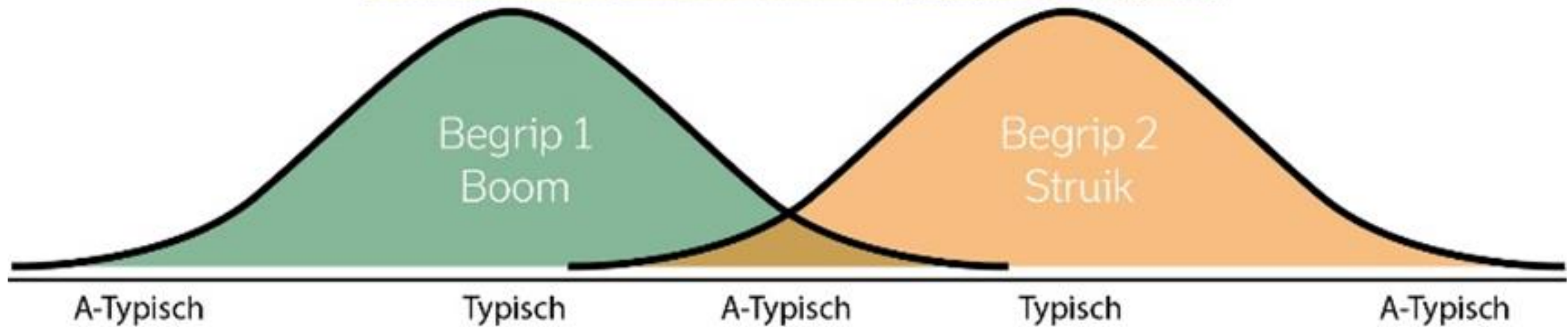
Boom: relatie kenmerken-individuen

	A	B	C	D	E	F	G	H	I	J	K
heeft boomvorm											
is houtachtig											
doet aan fotosynthese											
is een geheel											
heeft een enkele stam											
op land											
heeft wortels											
is ouder dan 10 jaar											
heeft takken											
heeft bladeren											
geen onderdeel van ander organisme											
% match	64	100	82	64	82	64	82	91	91	45	55

Boom: prototypemodel en exemplarenmodel



Boom of struik



Vraag aan ChatGPT: boom (incl. supertypen)

- Geef me de overeenkomsten en verschillen van definitie IMBOR-boom en IMGEO-boom: Een IMBOR-boom is een Een houtachtig gewas (loofboom of conifeer) met een wortelgestel en een enkele, stevige en houtige stam, die zich boven de grond vertakt en een Geo-object dat zich geheel materieel manifesteert en een Reëel object dat bestaat uit een aaneengesloten hoeveelheid vormvaste materie, primair bijeengehouden door interne krachten (zwaartekracht of elektromagnetische kracht) en een fenomeen in de werkelijkheid, dat direct of indirect is geassocieerd met een locatie relatief ten opzichte van de aarde en een Fysiek object (vormvast of niet-vormvast) dat in de werkelijkheid tastbaar en zichtbaar is (of kan zijn), door de mens gemaakt of natuurlijk ontstaan en een Object welke gegevens kent betreffende de vastlegging binnen registratieve ruimten en een Object welke door een organisatie beheerd wordt in de beheerfase van zijn levenscyclus. Een IMGEO-boom is een markante boom die geen onderdeel uitmaakt van een ander boom- of struikbeplanting en een Solitair vegetatieobject of lijn- of vlakvormige groep gelijksoortige vegetatieobjecten met een beperkte omvang en een grootschalig topografisch object, al dan niet uit de basisregistratie en een SolitaryVegetationObject en een CityObject.

Antwoord ChatGPT

Overeenkomsten:

1. Beide definities verwijzen naar bomen.
2. Beide definities impliceren dat de boom een afzonderlijk object is.

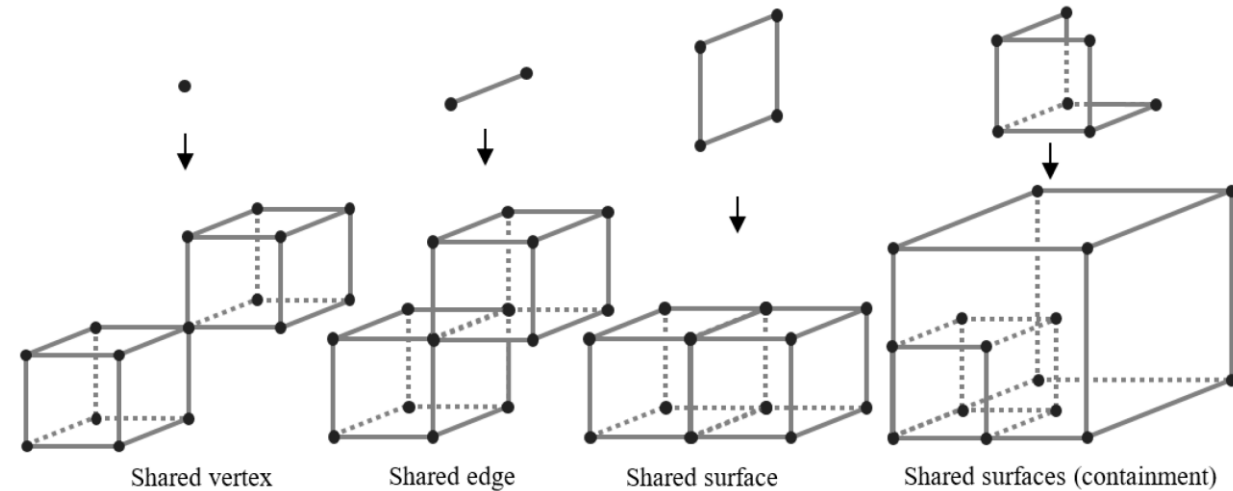
Antwoord ChatGPT (vervolg)

Verschillen:

- IMBOR-boom:
 1. De definitie van de IMBOR-boom geeft een gedetailleerde beschrijving van de fysieke kenmerken van een boom, zoals het zijn van een houtachtig gewas, de aanwezigheid van een wortelgestel en een enkele, stevige en houtige stam, en de vertakkingen boven de grond.
 2. De definitie benadrukt verschillende perspectieven van een object, zoals een geo-object, een reëel object, een fenomeen, een fysiek object, en een object dat gegevens bevat voor registratie en beheer.
- IMGEO-boom:
 1. De definitie van de IMGEO-boom legt de nadruk op het feit dat het een markante boom is die geen onderdeel uitmaakt van een andere boom- of struikbeplanting. Het benadrukt dat de boom op zichzelf staat en opvalt.
 2. De definitie verwijst naar verschillende categorisering van de boom, zoals een solitair vegetatieobject of een groep vegetatieobjecten met beperkte omvang. Het kan ook worden gezien als een grootschalig topografisch object of een CityObject, afhankelijk van de context.
- Over het algemeen ligt het verschil tussen de definities in de mate van detail en focus op fysieke kenmerken versus categorisering en contextuele aspecten van de boom. **De IMBOR-boom legt meer nadruk op de fysieke kenmerken en verschillende perspectieven van een boom, terwijl de IMGEO-boom meer nadruk legt op de markante eigenschap en categorisering van de boom binnen een ruimtelijke context.**

Geometrie: differentiatie naar:

- Topologie
- Geometrie
- Geografie



Geometrie: bevindingen

- Algemeen: per fysiek object: meerdere geometrieën (e.o.)
- IFC: net onderscheid topologie-geometrie-(relatieve) plaatsing
- CityGML: alle objecten absoluut geogerefereerd
- IFC: héél veel geometrische representaties
- CityGML: geometrie volgens ISO 19107: Point, Curve, Surface, Solid plus aggregaties zoals MultiPoint, MultiCurve en samenstellingen, zoals CompositCurve.
- Details: zie rapport.

Hoofdconclusies

1. Er zijn veel verschillen in de terminologie en de definities van de onderzochte objecttypen Gebouw, Weg, Kolk, Boom en Geometrie.
2. Semantische analyse leidt tot inzichten die wezenlijk zijn voor adequate gegevensintegratie.

Belangrijkste aanbevelingen

1. Werk T3D-begrippenmodel uit tot complete 3D ontologie; beter nog: tot een complete conceptuele stadsontologie
2. Verdiep en verbreed de uitgevoerde analyses.
3. Breid de query-, presentatie- en analysemogelijkheden uit, zo mogelijk tot een compleet instrumentarium voor metadataharmonisatie.

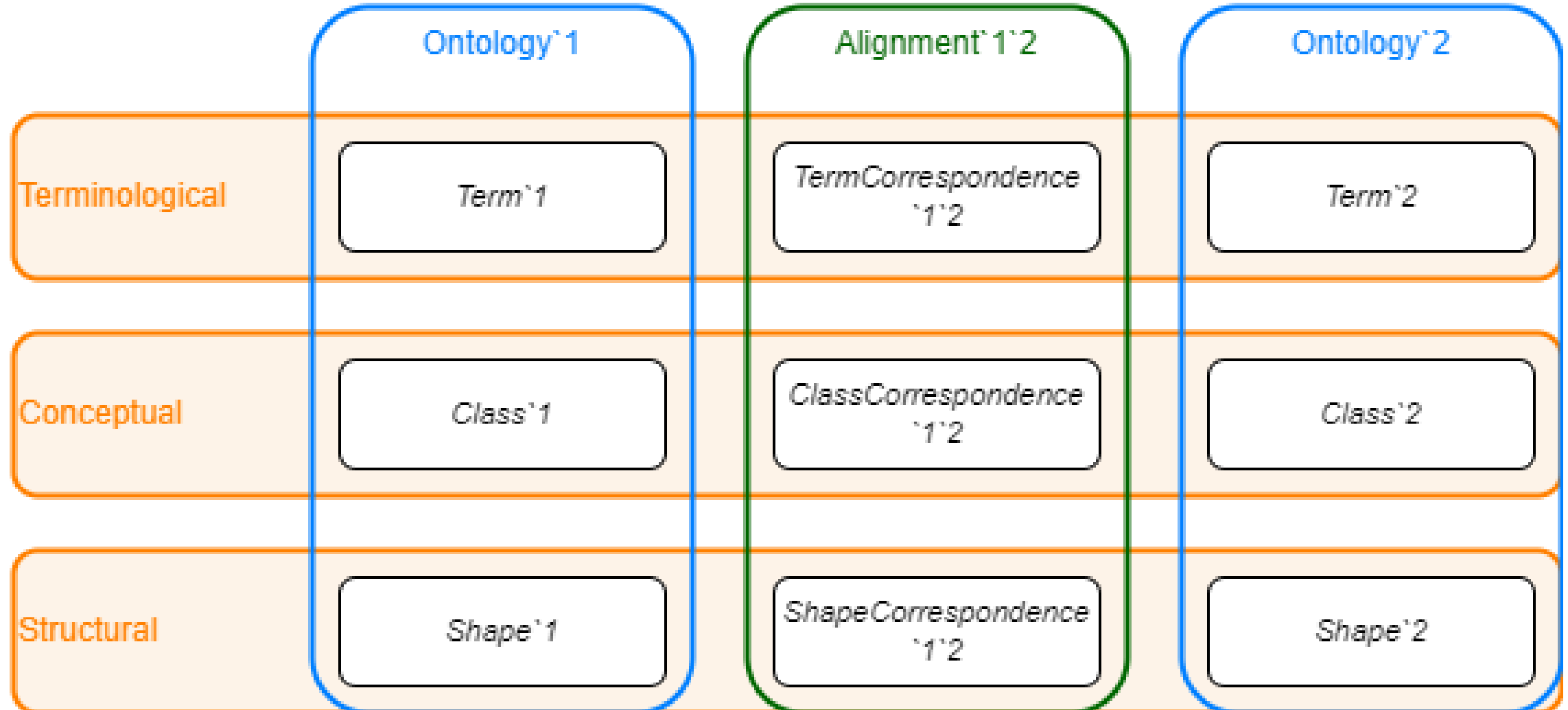
en.. denk nog eens na over de term ‘semantische analyse’: de analyse is breder en behelst ook terminologie, syntax en pragmatiek.
(denk aan de ‘semiotische driehoek’).

3. Twee extra's

3.1 Ontology alignment

3.2 Dataharmonisator

3.2 Ontology alignment



Voorbeeld: spuisluis



Spuisluis (ook wel 'Zijl')

- term-alignment:
Spuisluis (NEN 2767) → narrower term → Sluis (IMBOR)
- conceptuele alignment:
Spuisluis (NEN 2767) → subtype van → Sluis (IMBOR)
- structuur-alignment:
Spuisluis (NEN 2767) → subtype van → Sluis (IMBOR) met de voorwaarde
Sluis.sluisType=spuisluis

bron: CROW (Rik): ('Ontology Matching through alignment: a Best Practice', CROW whitepaper, draft document 09 February 2023')(vrij geïnterpreteerd door Hein C)

Matchingmethoden

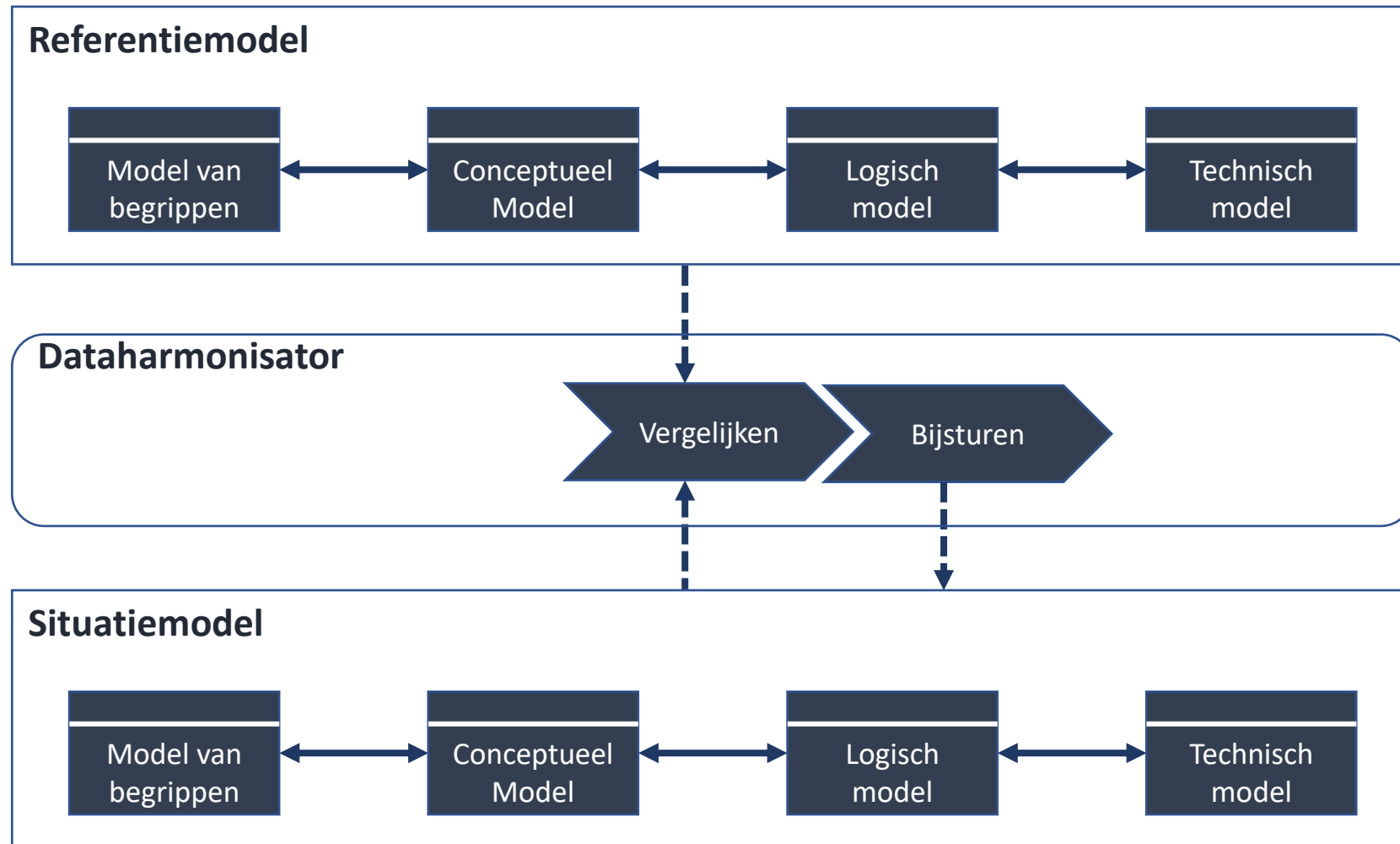
- elementniveau: string-based, linguistic-based, constraint-based en resource-based matching
- structureel niveau: taxonomy-based en instance-based matching
- expert-judgement

Zie bijvoorbeeld: Shenghui Wang (UT) : Ontology Matching, 22 March 2022

Voorbeeld:

Instance matching: Whether different instances refer to the same real-world entity in a given domain (e.g., a person, a place, a movie, a book, etc.)

3.1 (Meta)Gegevensharmonisator. Gemeentelijk Gegevenswoordenboek (GGw)



4. Terugkoppeling naar MIM

- Semantische analyse levert een bijdrage aan het doel van MIM: gemeenschappelijke afspraken voor het beschrijven en uitwisselen van gegevens
- Dit heeft primair betrekking op M1: Gegevensmodellen
- In het bijzonder op:
 - Conceptueel informatiemodel**
 - analyse en harmonisatie van concepten
 - Logisch gegevensmodel**
 - analyse en harmonisatie van klassen (objecttypen, attribuutsoorten en relatiesoorten) en constraints

5. Tot slot een drietal stellingen

1. Een semantische basis is een noodzakelijke voorwaarde voor een goede informatievoorziening.
2. Aandacht voor semantiek leidt tot inzicht voor beheerders en gebruikers in waar ze mee bezig zijn en daarmee tot een betere informatievoorziening.
3. De oplossing voor gegevensincongruentie ligt niet in een brute gelijkschakeling van definities, en ook niet in het oppervlakkig mappen op basis van verwante begrippen, maar in een zorgvuldig en langdurig proces van het afstemmen van ontologieën ('ontology alignment').

Die oplossing is overigens een punt op de horizon dat per definitie nooit bereikt kan worden.

Meer informatie?

Rapport T3D:

[Naar een ontologie voor 3D omgevingsgegevens](#), 6 april 2023

E-mail:

hein@corstens.nl