/// Verwerkingsdocument thematische werkgroep 1 OSLO Metadatastandaard voor services

Datum: 09/10/2020

Locatie: Teams meeting (virtueel)

INHOUD

Dit document bevat de verwerking van de eerste thematische werkgroep en dient ter voorbereiding van de tweede thematische werkgroep op 13/10/2020.

CONTEXT

Het initiatief voor dit standaardisatietraject komt vanuit Informatie Vlaanderen waarbij de doelstelling is om een semantisch model op te stellen voor het uitwisselen van metadata over diensten (services) dat is afgestemd met alle betrokken partijen. Dit traject zal de huidige OSLOstandaard DCAT-AP uitbreiden met de mogelijkheid om ook API's te beschrijven.

Het doel is tweeledig:

Als eerste willen we de metadata van services te standaardiseren vanuit het perspectief van de beheerder van een informatiesysteem in samenspraak met alle belanghebbenden. Vandaag wordt deze informatie in verschillende vormen en niveaus van kwaliteit voorzien, sterk afhankelijk van de gebruikte technologie en platform om deze datasets en services aan te bieden. Er is nood aan een overzicht.

Op basis van internationale standaarden (DCAT-AP v2.0) bouwen we het vocabularium op dat een uitbreiding zal zijn van de OSLO-standaard DCAT-AP met de mogelijkheid om ook API's te beschrijven.

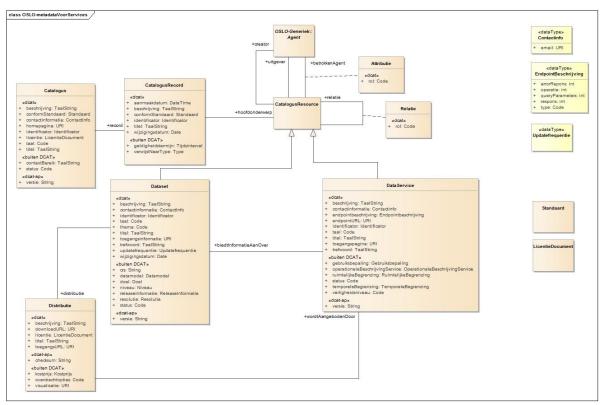
Ten tweede wordt er tegelijk gewerkt aan een referentie implementatie (piloot) voor het metadateren van services. Op deze wijze wordt reeds in een vroeg stadium de toetsbaarheid in het veld getoetst.

HOE HET VOORLOPIG SNEUVELMODEL LEZEN?

Het datamodel Metadatastandaard voor services beschrijft welke gegevens we voor het beschrijven van een service kunnen vastleggen, hoe deze gegevens gestructureerd zijn en wat de

verbanden tussen deze gegevens zijn. Het model bestaat uit **entiteiten**. Dit zijn de rechthoeken in het onderstaande model.

De specificiteit van entiteiten wordt uitgedrukt in eigenschappen. Die bestaan uit **attributen** en relaties. Een attribuut specificeert een bepaald kenmerk van de entiteit zonder naar een andere entiteit te linken; de attributen van een entiteit staan vermeld in de rechthoeken. Een **relatie** legt het verband tussen twee entiteiten; dit wordt aangeduid met een lijn. Wat de relatie inhoudt, staat er bij vermeld. In enkele uitzonderlijk gevallen zijn (ter vereenvoudiging van het model) relaties weergegeven als attributen.



Figuur 1: Eerste sneuvelmodel terug te vinden en beter te bekijken via link

Aangezien we ons model definiëren in de open wereld van *linked data*, betekent dit dat er nog eigenschappen zijn die toegevoegd kunnen worden, ook al staan deze niet expliciet vermeld in dit model. Zo, kunnen de entiteiten in ons model gekoppeld worden aan, o.a., de entiteiten in de modellen SCAT-AP Vlaanderen.

VOORLOPIG SNEUVELMODEL: DEFINITIES, ATTRIBUTEN EN RELATIES

Omdat er nood is aan een overzicht van informatie m.b.t de metadata van services vanuit verschillende perspectieven is het belangrijk dat we vertrekken vanuit de use cases die werden aangebracht tijdens de eerst business werkgroep. Door use cases te formuleren geven we een eerste invulling aan wat er mogelijk is met het ontwikkelde datamodel en komen we alvast tegemoet aan enkele behoeften.

Dit zijn enkele use cases die werden gedefinieerd in de business werkgroep van 3/09::

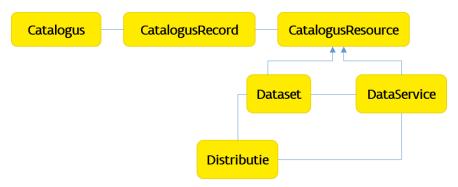
Vanuit het standpunt van een marktplaats:

- We willen een marktplaats waarop beschrijvingen van datasets, hun ontsluitingsvormen en services beschikbaar zijn.
- Een catalogus moet te importeren zijn in een andere catalogus.

Use cases vanuit de verschillende gebruikersgroepen:

- Als onderzoeker wil ik overheidsdata kunnen gebruiken in mijn analyse.
- Als ontwikkelaar wil ik een gemeentelijk beheersysteem koppelen aan de actuele data.
- Als ondernemer wil ik een publiek toegankelijke site bouwen met overheidsinformatie.
- Als overheid wil ik weten welk agentschap deze data beheert.
- Als marktplaats wil ik elk element weergeven/functionaliteiten mogelijk maken enkel op basis van de aangeleverde informatie in de catalogus.
- Als uitgever/beheerder van data wil ik dat potentiële afnemers mij eenvoudig vinden.
- Als uitgever/beheerder van data wil ik al mijn ontsluitingsvormen op een plek hebben.
- Als uitgever/beheerder van data wil ik dat het aanmaken van de beschrijvingen geïntegreerd kunnen worden in mijn ontwikkelomgeving.
- Als uitgever/beheerder van data wil ik de relaties met andere catalogusresources kunnen weergeven.
- Als uitgever/beheerder van data wil ik services bekend maken.

Daarnaast is het belangrijk dat vanuit elk perspectief (geografisch, REST, SOAP, Linked Data) wordt bekeken of het model de relevante kerneigenschappen bevat. Daarom vragen we de deelnemers om te kijken vanuit hun eigen perspectief en het sneuvelmodel toe te passen op een eigen service en te kijken of de eigenschappen al dan niet relevant zijn. Zo kunnen we samen onderstaande tabel opmaken die de basis zal vormen voor het model. Hieronder wordt een eerste sneuvelversie gepresenteerd.



Figuur 2: Vereenvoudigde visualisatie van het sneuvelmodel

Toelichting bij het tot stand komen van dit sneuvelmodel

Deze sneuvelversie kwam tot stand na de verwerking van de input uit de eerste thematische werkgroep waarbij er door de deelnemers verschillende voorstellen voor definities en voorstellen voor eigenschappen werden aangebracht. Deze werden verwerkt en waar mogelijk gemapt op reeds bestaande termen uit de internationale standaard DCAT-AP v2.0 (waarmee we ook mee willen aligneren). De eigenschappen die nog niet werden teruggevonden in internationale standaarden staan aangeduid in het blauw omdat we hier nog geen definitie voor gevonden hebben. De eigenschappen in het oranje zijn ook niet opgenomen in DCAT maar hier bestaat wel een mapping voor.

OEFENING

Bekijk onderstaande tabel met voorgestelde definities, attributen en relaties en probeer vanuit je eigen perspectief en vanuit de use cases na te denken of jouw service hierin beschreven kan worden.

Term	Definitie	
DataService	<u>Definitie</u>	
	Een verzameling van operaties die ofwel toegang tot één of meer	
	datasets voorzien ofwel	een dataverwerkingsdoel hebben.
	Attributen	
	beschrijving	Descriptie van het item.
	contactInformatie	Relevante contactinformatie voor de dataservice.
	endpointbeschrijving	Een beschrijving van de diensten die beschikbaar zijn via de end-points, met inbegrip van hun werking, parameters, enz.
		Of machineleesbare beschrijving of een tekstuele pagina?
	endpointURL	De rootlocatie of het primaire eindpunt van de dienst (een webresolvable URI).
	identificator	Een unieke identificator van het item.

	-
taal	Een taal van het item. Dit verwijst naar de natuurlijke taal die wordt gebruikt voor tekstuele metagegevens (d.w.z. titels, beschrijvingen, enz.) van een catalogus of de tekstuele waarden van een datasetdistributie.
titel	Een naam die aan het item wordt gegeven.
toegangspagina	Een webpagina waarnaar een webbrowser kan worden genavigeerd om toegang te krijgen tot de catalogus, een dataset, de distributie en/of aanvullende informatie.
trefwoord	Een trefwoord of tag die de resource beschrijft.
gebruiksbepaling	
operationeleBeschrij vingService	
ruimtelijkeBegrenzin g	
status	
temporeleBegrenzin g	
veiligheidsniveau	
versie	

Volgende eigenschappen werden vermeld tijdens de eerste thematische werkgroep maar zijn voorlopig nog niet opgenomen omdat ze eerder een groep van informatie bevatten dan een specifieke eigenschap:

- Authenticatie en toegangsinformatie
- Wettelijke beperkingen
- Extent van de dataservice
- Quality of service
- Te gebruiken tot/afneembaar tot
- Bron
- Data elementen
- Distributieprotocollen
- Gebruikte/extra WS afhankelijkheden
- Actuele/historische gegevens
- Doelpopulatie
- Hydra-cg (hypermedia gebruik)

De werkgroep wordt gevraagd om suggesties/usecase beschrijvingen te geven de invulling van deze informatie in de vorm van concrete eigenschappen opdat ze eventueel kunnen worden opgenomen. Relaties Welke bijkomende relaties tussen "Dataservices" moeten worden opgenomen? Catalogus <u>Definitie</u> Een beheerde verzameling van informatie over resources. <u>Attributen</u> beschrijving Descriptie van het item. conformStandaard Een vastgestelde standaard waaraan de beschreven resource voldoet. contactInformatie Relevante contactinformatie voor de catalogus. homepagina Een homepagina van de catalogus (een publiek webdocument meestal beschikbaar in HTML). identificator Een unieke identificator van het item. licentie Een juridisch document op grond waarvan de resource te beschikking wordt gesteld. taal Een taal van het item. Dit verwijst naar de natuurlijke taal die wordt gebruikt voor tekstuele metagegevens (d.w.z. titels, beschrijvingen, enz.) van een catalogus of de tekstuele waarden van een datasetdistributie. titel Een naam die aan het item wordt

gegeven.

contextBereik

status

versie

Volgende eigenschappen werden vermeld tijdens de eerste thematische werkgroep maar zijn voorlopig nog niet opgenomen omdat ze eerder een groep van informatie bevatten dan een specifieke eigenschap:

- Wettelijke aspecten
- Distributie mogelijkheden
- Context/Bereik
- Extent van de data

De werkgroep wordt gevraagd om suggesties/usecase beschrijvingen te geven de invulling van deze informatie in de vorm van concrete eigenschappen opdat ze eventueel kunnen worden opgenomen.

Relaties

gebruikt	Een omschrijving van een begrip.

Welke bijkomende relaties tussen "Catalogi" moeten worden opgenomen?

CatalogusRecord

Definitie

Een record in een catalogus dat de registratie van een catalogus resource beschrijft.

Attributen

1001100110011	
aanmaakdatum	De datum van opname (d.w.z. de formele registratie) van de overeenkomstige dataset of dienst in de catalogus.
beschrijving	Descriptie van het item.
conformStandaard	Een vastgestelde standaard waaraan de beschreven resource voldoet.
identificator	Een unieke identificator van het item.
titel	Een naam die aan het item wordt gegeven.
wijzigingsdatum	Meest recente datum waarop de catalogusvermelding werd gewijzigd, bijgewerkt of aangepast.

hoofdonderwerp	De dcat:Recource (dataset of service) beschrijven in het record.
geldigheidstermijn	
verwijstNaar	
versie	
Relaties	

verwijstNaarType

Dataset

<u>Definitie</u>

Een door één agent beheerde of gepubliceerde verzameling van data, en toegankelijk in één of meerdere voorstellingswijzen.

<u>Attributen</u>

Attributeri	
beschrijving	Descriptie van het item.
contactinformatie	Relevante contactinformatie voor de dataset.
identificator	Een unieke identificator van het item.
taal	Een taal van het item. Dit verwijst naar de natuurlijke taal die wordt gebruikt voor tekstuele metagegevens (d.w.z. titels, beschrijvingen, enz.) van een catalogus of de tekstuele waarden van een datasetdistributie.
thema	Een hoofdcategorie van de resource. Een resource kan meerdere thema's hebben.
titel	Een naam die aan het item wordt gegeven.
toegangsInformatie	
trefwoord	Een trefwoord of tag die de resource beschrijft.
updatefrequentie	De frequentie waarmee de dataset wordt gepubliceerd.
wijzigingsdatum	Meest recente datum waarop de catalogusvermelding werd gewijzigd, bijgewerkt of aangepast.

Volgende eigenschappen werden vermeld tijdens de eerste thematische werkgroep maar zijn voorlopig nog niet opgenomen omdat ze eerder een groep van informatie bevatten dan een specifieke eigenschap:

- Inwinning van de data
- Kwaliteit van de data
- Wettelijke beperkingen
- Extent
- Distributiemogelijkheden
- Verwijzing naar het datamodel/data elementen
- Doel van gebruik
- Bewerkingen (lineage)
- Applicatieschema

De werkgroep wordt gevraagd om suggesties/usecase beschrijvingen te geven de invulling van deze informatie in de vorm van concrete eigenschappen opdat ze eventueel kunnen worden opgenomen.

Welke bijkomende relaties tussen "Datasets" moeten worden opgenomen?

Distributie

Definitie

Een specifieke voorstelling van een dataset, gekenmerkt door een of meerdere eigenschappen zoals taal, format, temporele en ruimtelijke dekking, granulariteit, ...

Attributen

beschrijving	Descriptie van het item.
downloadURL	De URL van het te downloaden bestand in een bepaald formaat. Bijv. CSV-bestand of RDF-bestand. Het formaat wordt aangegeven door dct:format en/of dcat:mediaType van de distributie.

licentie	Een juridisch document op grond waarvan de resource te beschikking wordt gesteld.
titel	Een naam die aan het item wordt
	gegeven.
toegangsURL	Een URL van de resource die toegang
	geeft tot een distributie van een
	dataset. Bijv. Landingspagina, feed,
	SPARQL endpoint.
ah a akayya	
checksum	
kostprijs	
overdrachtOpties	
visualisatie	

Volgende eigenschappen werden vermeld tijdens de eerste thematische werkgroep maar zijn voorlopig nog niet opgenomen omdat ze eerder een groep van informatie bevatten dan een specifieke eigenschap:

• Link naar de applicatie waar de data kan bekeken worden (nu in het model als attribuut "Visualisatie")

De werkgroep wordt gevraagd om suggesties/usecase beschrijvingen te geven de invulling van deze informatie in de vorm van concrete eigenschappen opdat ze eventueel kunnen worden opgenomen.

Relaties

wordtAangebodenD	Een dataservice die toegang geeft
oor	tot de distributie van de dataset.

CatalogusResource

Definitie

Een resource gepubliceerd of beheerd door 1 agent.

<u>Attributen</u>

/

Relaties

betrokken agent	
creator	De entiteit die verantwoordelijk is voor het aanmaken van de resource.

relatie		
uitgever	De entiteit die verantwoordelijk is voor de resource beschikbaar te maken.	

EXTRA OPMERKINGEN EN VRAGEN BIJ HET MODEL:

Mapping op het ISO meta-metadata model

De mapping van het ISO meta-metadata model met deze datastandaard is een onderdeel van de POC.

Merk op dat de doelstelling hier niet is om alle verplichtingen van ISO meta-metadata over te nemen als verplichtingen voor andere domeinen. Dataservices die onder de verplichting vallen om INSPIRE compliant metadata aan te leveren moeten de hiervoor betreffende afspraken volgen.

• Relaties tussen Datasets, Distributies, Dataservices en Catalogi beschrijven.

Welke bijkomende relaties zijn nodig tussen de verschillende entiteiten?

Bijvoorbeeld tussen Distributie en Dataservice?

• Bekijk het model ook vanuit het perspectief FileServer

Welke eigenschappen zijn belangrijk voor een dataservice dat een FileServer aanbiedt? Zijn hier specifieke eigenschappen die nog niet opgenomen zijn in dit sneuvelmodel?

• Interactie met een Agent

Welke rollen zijn belangrijk? De uitgever, beheerder, ..? Hoe zit dat in relatie met de contactinformatie?

FEEDBACK GEVEN OP DIT VERWERKINGSDOCUMENT

Er zijn twee mogelijkheden om feedback te geven:

- Via mail naar informatie.vlaanderen@vlaanderen.be
- Vermeld het OSLO-thema: metadataVoorServices

of

- Via GitHub https://github.com/Informatievlaanderen/OSLO-Standaarden/tree/master/standaarden-in-ontwikkeling/metadata-voor-services

Deze feedback zal meegenomen in de volgende werkgroep van 13 oktober.

AANKOMENDE WERKGROEPEN

Indien u (of een collega) graag zou willen deelnemen aan een van de aankomende werkgroepen, kan u via de onderstaande link een overzicht van de workshops terugvinden en u ook zo inschrijven. De tweede thematische werkgroep is de eerstvolgende werkgroep en zal plaatsvinden op dinsdag 13 oktober 2020 om 9u via Microsoft Teams waarvan de link wordt doorgestuurd naar de deelnemers.

Inschrijven kan via https://overheid.vlaanderen.be/opleiding/publieke-werkgroepen-oslo-metadatastandaard-voor-services