**///** OSLO LDES: Business Werkgroep

**////////////////////////////////////////////////////////////////////////////////////////////////////////**

Datum: 16/12/2021

Locatie:  Teams meeting (virtueel)

Moderators: Thijs Hegge, Jitse De Cock, Pieter Colpaert, Annelies De Craene, Geert Thijs, Laurens Vercauteren, Arne Van Der Stuyft

**////////////////////////////////////////////////////////////////////////////////////////////////////////**

**Aanwezigen**

* Digitaal Vlaanderen
  + Thijs Hegge
  + Annelies De Craene
  + Geert Thijs
  + Laurens Vercauteren
  + Arne Van Der Stuyft
  + Jitse De Cock
  + Veerle Beyaert
  + Filip Borloo
  + Liesbet De Wolf
  + Michael Mampaey
  + Sarah Smits
  + David Van den Brande
  + Dwight Van Lancker
  + Geert Van Nuffelen
  + Frederik Schodts
  + Simon Claus
  + Veerle Beyaert
  + Adriaan Walpot
  + Geraldine Nolf
* IMEC
  + Pieter Colpaert
  + Philippe Michiels
  + Birger Schrevens
* Departement Omgeving
  + Geert Van Haute
* IDLab - Universiteit Gent
  + Ben De Meester
  + Brecht Van de Vyvere
  + Wout Slabbinck
  + Els de Vleesschauwer
  + Melanie Verstraete
* District09
  + Geert Limpens
  + Maarten Segers
* Vlaams Instituut Voor de Zee (VLIZ)
  + Britt Lonneville
  + Milan Lamote
  + Marc Portier
* Maritieme Dienstverlening en Kust
  + Tim Vereecken
* De Watergroep
  + Pieter Jan Haest
* Agentschap Wegen en Verkeer
  + Joris Hoogeboom
* Smart Region Limburg
  + Geert Thoelen
* Atos
  + Marco Tiberius
  + Rudy De Smedt
  + Louis Oberije
  + Casper Gehner
  + Bert Wissink
* Cegeka
  + Mitchel Smedts
  + Sam Van der Cammen
* Agoria
  + Jelle Hoedemaekers
* Integration Architect Archers
  + Joost Vandebosch
* Konsolidate
  + Christophe Cop
* Esri Belux
  + Gert Bergers
  + Wim Driessens
* Movias
  + Steven Logghe
* MyCSN NV
  + Jan Geukens
* Redpencil.io
  + Jo Delaure
* Ziekenhuis Oost-Limburg
  + Noella Pierlet

**Agenda van de werkgroep**

|  |  |
| --- | --- |
| 09u00 – 09u10 | **Welkom** |
| 09u10 - 09u30 | **Context & Linked Data Event Streams (LDES)** |
| 09u30 – 09u40 | **Open Standards for Linked Organisations (OSLO)** |
| 09u40 – 10u10 | **Inspiratie (Bestaande standaarden en use cases)** |
| 10u30 – 10u40 | **Pauze** |
| 10u40 – 11u50 | **Brainstorm oefeningen** |
| 11u50 – 12u | **Volgende stappen** |

1. **Inleiding**

Het doel van dit traject is om de structuur van Linked Data Event Streams (LDES) te standaardiseren in samenspraak met de belanghebbenden. Dit zal het mogelijk maken om datastromen semantisch te verrijken, wat synchronisatie met, en replicatie van de LDES-brondata vereenvoudigt alsook beter combineerbaar maakt met je eigen Linked Data Event Stream die op de LDES-brondata is gebaseerd.

Op basis van internationale standaarden bouwen we een vocabularium en een applicatieprofiel op. De OSLO LDES standaard vormt mee de basis voor het uitbouwen van het relance project ‘Vlaamse Sensor Data Space’.

De standaardisering gebeurt op vraag van de verschillende belanghebbenden binnen de sector, met ondersteuning van Digitaal Vlaanderen.

* 1. **Context Linked Data Event Stream**

*[We verwijzen naar de slides voor meer informatie.]*

* 1. **Context OSLO:** **Open Standaarden voor Linkende Organisaties**

*[We verwijzen naar de slides voor meer informatie.]*

De OSLO-context wordt meegegeven, en meer specifiek het proces en de methodes. De tijdlijn van het project wordt eveneens overlopen.

De Vlaamse overheid zet in op eenduidige standaarden voor de uitwisseling van informatie. Het is de bedoeling om zo te zorgen voor meer samenhang en een betere vindbaarheid van data. Op die manier kan iedereen de gegevens makkelijker gebruiken. Met OSLO wordt er concreet ingezet op semantische en technische interoperabiliteit. De vocabularia en applicatieprofielen worden ontwikkeld in co-creatie met Vlaamse administraties, lokale besturen, federale partners, de Europese Commissie en private partners (ondertussen meer dan 400 bijdragers).

Momenteel zijn er reeds 99 erkende standaarden, 40 kandidaat standaarden en 18 standaarden in ontwikkeling. De standaard voor LDES behoort tot deze laatste categorie. Meer informatie over OSLO kan hier teruggevonden worden: <https://overheid.vlaanderen.be/oslo-wat-is-oslo> en <https://data.vlaanderen.be/>.

1. **Brainstormsessie: identificeren van informatie**

De doelen en de aanpak van de virtuele brainstormsessie worden hieronder beschreven. Tevens worden de belangrijkste bevindingen uit deze werkgroep samengevat.

* 1. **Doel van de werkgroep**
* Toelichten van de context en inleiding over OSLO
* Informatienoden en use cases capteren
* Prioriteren van de informatienoden van deze OSLO-standaard
  1. **Aanpak**

*[We verwijzen naar de slides voor meer informatie.]*

De sessie werd opgesplitst in de volgende delen:

* Identificeren van de verwachtingen van de deelnemers aangaande deze werkgroep van, meer algemeen, dit OSLO LDES traject;
* Oplijsten van de domein-specifieke problemen/uitdagingen die de deelnemers opgelost wensen te zien m.b.v. een LDES;
* Identificeren van de zaken die de deelnemers nodig achten om hun problemen/uitdagingen op te lossen;
* Uit de geïdentificeerde benodigdheden de terminologie extraheren die volgens de deelnemers mee opgenomen dient te worden in de specificatie; en
* Identificeren van bestaande datamodellen, informatiemodellen of standaarden die mogelijks affiniteit hebben met het traject of die mogelijks hergebruikt zouden kunnen worden.
  1. **Belangrijkste informatienoden en informatie-elementen**

De deelnemers gaven, a.d.h.v. post-its, aan wat zij de belangrijkste informatienoden vonden, via een brainstorming tool (Mural).

In de Github-repository van dit traject, kan u steeds issues (dit mogen opmerkingen, vragen of aanbevelingen zijn) aanmaken. Het overzicht van deze issues kan geraadpleegd worden via onderstaande link: <https://github.com/Informatievlaanderen/OSLOthema-ldes/issues>

* + 1. **Verwachting van het OSLO LDES traject**

In de eerste oefening werd aan de deelnemers gevraagd om hun verwachtingen van de business werkgroep en het traject op te lijsten.De verwachtingen van het traject zijn opgedeeld in zelf-gedefinieerde clusters waar mogelijk, dewelke zijn aangemaakt na de werkgroep door de organisatoren:

**Informatie inwinnen over LDES:**

* Kennis opdoen over een LDES en de mogelijkheden ervan ontdekken.
* Te weten komen wat de minimale requirements zijn voor het gebruik van een LDES.
* Een heldere scope en governance van een standaardisatieproces verkrijgen.
* Kennis opdoen over de plannen van de Vlaamse overheid om Linked Open Data (LOD) vorm te geven.
* Evalueren of een LDES tegemoet kan komen aan de vereisten in een project rond data aggregatie en redistributie.
* Betrokkenheid van private dataleveranciers bekomen.
* Achterhalen of een LDES ook gebruikt kan worden voor niet-open linked data.
* Inzicht in gebruikersrechten en schaalbaarheid van de standaard zowel in Vlaanderen (volledigheid) als internationaal verkrijgen.
* Op de hoogte te blijven van initiatieven rond data.
* Weten wat de meest kostenefficiënte interface design is voor een LDES (serialisatie formaat, paginagrootte).
* Te weten komen of/hoe LDES te combineren is met een door wetgeving opgelegd dataformaat.
* Op zoek naar een platform waar tools (validator, consumer, docker containers, indexers, produce libs and patterns ...) en ervaringen rond deze standaard kunnen uitgewisseld worden.
* Te weten komen op welke manier er geconnecteerd wordt aan Europese en internationale standaarden.
* Het in productie nemen van deze specificatie en dus ook rekening houden met bestaande data, services, publicaties, metadata, versiebeheer, ... voor het gebruiksgemak, beheersgemak, …
* Een gedragen definitie van de elementen in mijn data stroom (~OSLO).

**Standaard voor het aanbieden van data, ongeacht de context:**

* OSLO LDES als een breed gedragen standaard om use cases te kunnen faciliteren.
* Een LDES dient makkelijk automatisch opgepikt te worden door platformen die optreden als broker voor de vindbaarheid (register + metadata van domeinen).
* Een gedragen specificatie waarmee de datadeling voor fast en slow moving data voor het landschap van overheden en bedrijven in Vlaanderen de drempels mee kan helpen wegwerken.
* OSLO data als LDES kunnen ontsluiten.
* Via een LDES de DCAT entries op het platform van VLIZ kunnen publiceren.
* Aanlevering van data via LDES onderzoeken.
* De Vlaamse overheid GIS-gerelateerd ondersteunen.
* Meer data beschikbaar laten maken om te kunnen innoveren.
* Een uniforme toegang tot sensor data.

**Metadata van een LDES**

* Bij de creatie van LDES ook metadata vol-automatisch wordt opgemaakt en gevoed, liefst automatisch.
* Een metadata van LDES-datasets en LDES-services kunnen ophalen, zodat ze in de Informatiecatalogus en Datavindplaats te vinden zijn.

**Synchronisatie en replicatie van data:**

* Een standaard om data replicatie problemen op te lossen.

**Overige:**

* Een LDES goed kunnen integreren in het SOLID verhaal.
* Het ArcGIS software platform een plaats te geven en integraal deel te laten uitmaken van het OSLO LDES initiatief/ecosysteem.
* Een referentie-implementatie zien op verschillende soorten data en services (geo, open, gesloten), die deze specificatie toetst tegen de praktijk.
* OSLO implementatiemodellen ondersteunen richting LDES en niet alleen applicatieprofielen.
* Een goed overzicht van master data systemen en data informatie flows en welke standaarden hiervoor van toepassing zullen zijn.
  + 1. **Welke problemen/uitdagingen zouden opgelost/verlicht kunnen worden m.b.v. een LDES?**

Als vervolgoefening op de verwachtingen die de deelnemers hebben van dit traject werd aan hen gevraagd welke problemen/uitdagingen d.m.v. een LDES verlicht zouden kunnen worden. Gerelateerd hieraan werd er aan hen gevraagd om de specifieke zaken die ze hiervoor nodig hebben via een post-it in Mural mee te delen.

De problemen/uitdagingen die de deelnemers ervaren zijn:

**Standaard voor het aanbieden van data:**

* Een gebrek aan standaardisatie en interoperabiliteit.
* Nood aan standaardisatie van (sensor)data.
* Geen uniforme mogelijkheid tot het aansluiten op een (Europese) dataspaces.
* Geen uniforme manier van werken voor tal van business use cases.
* Behoefte aan data-uitwisseling over verschillende sectoren heen.
* Datastromen niet commercieel kunnen aanbieden.

**Replicatie en synchronisatie**

* Het steeds moeten wachten op data dumps/volledige datasets.
* Geen synchronisatie van geactualiseerde data.
* Geen (automatische) data-verrijking.

**Versiebeheer**

* Geen versiebeheer.
* Moeilijk om bron en historiek van een fast-moving dataset te kunnen vinden.

**Maintenance hell**

* Formatting hell/veelheid aan gevraagde te ondersteunen query-API’s.
* Geen schaalbaarheid om rechtstreeks applicaties te laten draaien op bestaande API’s.
* Legacy toepassingen.

**Overige**

* Te lage kwaliteit van brondata.
* Complexe technologie & te weinig developers met kennis van het Semantische Web.
* Geen correcte categorisatie van de metadata (bij fast-moving data).

Bij het oplijsten van wat de deelnemers nodig hebben om hun problemen/uitdagingen opgelost of verlicht te zien worden, kwamen volgende zaken naar boven:

* Versiebeheer.
* Retentiebeleid.
* Opleidingen om een LDES te kunnen implementeren.
* Use cases/voorbeelden van een LDES in de praktijk.
* Integratiepatronen.
* Een gedragen en goed gedocumenteerd applicatieprofiel.
* Vocabularium.
* Gebruiksgemak aan de ontvanger zijde (gebruikers buy-in)
* LDES plugins.
* Connector die data van format x naar format y vertaalt.
* Intermediates (die LDES feeds omzetten in specifieke fragments, of voeding van bepaalde query API).
* Uniforme library waarin de context van alle objecten wordt beschreven.
* Basisregisters.
* Optionele toegangscontrole.
* Community (forum).
* Een LDES standaard die overweg kan met implementatiemodellen.
* Niet-technische aspecten (privacy, security, kosten-baten, …)
* Mogelijkheid tot behoud van bestaande standaarden.
* Throttling kunnen beheren.
* Mogelijkheid tot opvragen van historische data.
* Levensloop van een dataset.
* Eenvoudige semantische toegang & uplifting.
* Harmonisatie met voldoende technische flexibiliteit.
* Requirements van een LDES.
* Duiding wanneer men wel/niet een LDES kan gebruiken.
  + 1. **Terminologie die (volgens de deelnemers) mee opgenomen dient te worden.**

Als vervolg op oefening 2.3.1 en 2.3.2. werd er aan de deelnemers gevraagd om de terminologie die zij in deze standaard wensen opgenomen te zien op te lijsten.

* Link naar de metadata-catalogus
* Retentiebeleid
* Versiebeheer
* Basisdefinitie LDES
  + Collectie
  + Members
  + Shape
  + Basis-LDES
* Afgeleide datasets
  + Data provenance / lineage
  + Versie-materialisatie
* Best practices/voorbeelden van een LDES
* Push-pull systeem
  + Vanaf >10s acceptable latency heeft een datagebruiker baat bij websockets/webhooks (push-based systeem), Push wordt wel ondersteund
  + Voor heel veel kleine bronnen (bv. SOLID-pods) moet die ook mogelijk zijn!
  + Resource-based werken is goed voor de vindbaarheid.
  + De bron moet weten waar de laatste versie zich bevindt zodanig dat op een termijn kan bevraagd (pull/push) kan worden.
* Geografie-gerelateerde data specificaties
  + Bestaande standaarden die er zijn mogen niet over het hoofd gezien worden en de koppeling met de specificaties (WMS, WFS, WMFS, …) moeten meegenomen worden.
* Gebruiksrechten data licentie
* Verhouding tot dcat:dataset of dcat:dataservice
* OGC-API features
* Informatie over de sensor
  + Niet elke sensor heeft dezelfde kwaliteit. Dit zou een element moeten zijn van het datamodel. Meetfouten, spreiding, kwartielen… Dit ter interpretatie van de gegevens.
    1. **Bestaande standaarden en modellen die kunnen hergebruikt worden**

De laatste oefening van de business werkgroep was het identificeren van bestaande datamodellen, informatiemodellen of standaarden die mogelijks affiniteit hebben met het traject of die mogelijks hergebruikt zouden kunnen worden. Hieronder wordt de verzamelde input tijdens de werkgroep weergegeven.

Binnen deze oefening werd er een onderscheid gemaakt tussen de standaarden waarop het datamodel zich effectief kan baseren en standaarden die de deelnemer relevant achten voor dit datamodel.

**Standaarden waarop het datamodel zich kan baseren:**

* NEN3610
* Fedora specificatie
* API-standaard
* OAI-PMH
* RSS/ATOM bij Wegen en Verkeer
* OGC webservices voor replicatie en synchronisatie
* Linked Data Platform 1.0.
* ABB LBLOD delta service
* Voor metadata OSLO Standaarden
  + Metadata - DCAT
  + DCAT - AP VL
  + GeoDCAT - AP (in opbouw)
* Cloudevents specificatie

**Standaarden die de deelnemers relevant achten in het datamodel:**

* Bestandsformaten
  + JSON
  + GeoJSON
  + CSV
  + GeoRSS
  + GML
  + …
* ISO 19110
* OSLO Applicatieprofielen

1. **Volgende stappen**

Indien u graag zou willen deelnemen aan één van de aankomende werkgroepen, kan u via de volgende [link](https://overheid.vlaanderen.be/opleiding/oslo-ldes) een overzicht van de workshops terugvinden en u ook zo inschrijven. De ‘thematische werkgroep 1 OSLO LDES’ is de eerstvolgende werkgroep voor het OSLO-traject en zal plaatsvinden op 20 januari 2022 om 9u00 via Microsoft Teams waarvan de link wordt doorgestuurd naar de deelnemers.

Inschrijven voor de volgende werkgroep (20 januari 2022) kan via onderstaande link: <https://overheid.vlaanderen.be/informatie-vlaanderen/agenda/thematische-werkgroep-1-oslo-ldes>