**///** OSLO Waterkwaliteit: Tweede thematische werkgroep

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

Datum: 01/02/2022

Locatie:  Teams meeting (virtueel)

Moderators: Maxime Pittomvils, Geert Thijs, Arne Van Der Stuyft

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

**Aanwezigen**

* Digitaal Vlaanderen
  + Maxime Pittomvils
  + Geert Thijs
  + Arne Van Der Stuyft
* Vlaamse Milieumaatschappij
  + Frank Lavens
  + Pieter Borremans
  + Greet Devriese
  + Jurgen Meirlaen
  + Dirk Goossens
  + Jonas Rooman
* Departement Omgeving
  + Tom Boonen
  + Geert Van Haute
  + Annelies Baert
  + Steven Geirnaert
* Provincie Antwerpen
  + Donald Vergauwe

**Agenda van de werkgroep**

|  |  |
| --- | --- |
| 13u00 – 13u10 | **Welkom en introductie** |
| 13u10 - 13u20 | **Samenvatting: wat hebben we gedaan in de vorige werkgroep?** |
| 13u20 - 15u00 | **Nieuwe aanpassingen aan het model**   * **Emissie(bron)** * **HydromorfologischeIndexType** * **Meetpunt en Meetnet** * **Aggregaties: somparameters/kenmerken en statistiek** * **Kwaliteit** |
| 15u00 - 15u50 | **Toelichtingen model**   * **Kenmerktype & chemische observaties** * **Beoordeling van waterkwaliteit** * **ChemischAgensVracht voorbeeld** |
| 15u50 – 16u00 | **Volgende stappen**   * **Publieke review** * **Vervolgtraject codelijsten** |

1. **Inleiding**

Het doel van dit traject is om de structuur rond waterkwaliteit te standaardiseren in samenspraak met alle belanghebbenden. Dit zal het mogelijk maken om datastromen semantisch te verrijken, data over waterkwaliteit beter vindbaar en begrijpbaar te maken, alsook beter combineerbaar met andere databronnen.

Op basis van het Europees model [ODALA Air & Water](https://purl.eu/ns#Applicationprofile) bouwen we onze standaard op voor Waterkwaliteit en brengen we dit in lijn met de geldende Vlaamse regelgeving.

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

* 1. **Samenvatting van de eerste thematische werkgroep**

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

De stappen die werden genomen in de eerste thematische werkgroep werden overlopen (zie slide 9). Tijdens de eerste thematische werkgroep werden de stappen uit de business werkgroep overlopen, vervolgens werd de nieuwe iteratie van het datamodel gepresenteerd waarop feedback werd verzameld. Tot slot werd er over enkele concepten (bv. uniformisatie van benamingen) gediscussieerd of deze binnen of buiten de scope van dit OSLO Waterkwaliteit-traject liggen.

* 1. **Scope van het project**

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

Het project is gericht op het ontwikkelen van een semantisch framework voor het in kaart brengen van de waterkwaliteit en het delen van data hieromtrent. Zo kan er een duurzaam vocabularium en applicatieprofiel (~datamodel) worden opgebouwd waarbij de focus ligt op de uitwisseling van data over waterkwaliteit.

Daarnaast werd er vermeld dat we voor het OSLO Waterkwaliteit model zijn vertrokken van [ODALA Air & Water - Waterquality](https://purl.eu/doc/applicationprofile/AirAndWater/Water/) model. Hier werd niet verder op ingegaan.

1. **Nieuwe iteratie van het datamodel**

De basiswerking werd aan de hand van het datamodel en een kort voorbeeld toegelicht. Hierop stelde de deelnemers volgende vragen:

* Is de instantiatie van een subklasse en een superklasse technologie specifiek? bv. In RDF neemt de superklasse als attributen van de subklassen over.

*In UML termen is dat niet de bedoeling. Er is een mismatch tussen UML en rdf. Een superklasse erft in UML niet de attributen van diens subklasse over.*

*Er werd gevraagd om deze vraag op te geven als issue om dit verder te kunnen onderzoeken. Tijdens het testen van voorbeelddata door de deelnemers op basis van dit model zal het duidelijk worden hoe de data concreet gerepresenteerd wordt in RDF.*

* Kan je verschillende types van waterkwaliteit observaties gelijktijdig meten?

*Dit is mogelijk en wordt in het model aangegeven doordat er ‘disjoint’ niet aanwezig is in het model. Op deze manier is double typing (twee verschillende types van waterkwaliteit meten) mogelijk. Disjoint zal expliciet toegevoegd worden aan het model.*

*De zes verschillende types van waterkwaliteit observaties die weergegeven zijn in het model zijn overigens niet alle mogelijke types. De gebruikers van het model kunnen er nog andere aan toevoegen tijdens de implementatie.*

* Waarvoor staat Bemonsteringsobject juist?

*Een ruimtelijk bemonsteringsobject is een plek waar iets geobserveerd wordt. Een bemonsteringsobject is wat je gaat observeren. Beide komen uit ISO Observaties & Bemonstering.*

* 1. **Aanpassingen aan het datamodel**

[We verwijzen naar de slides voor meer informatie]

* + 1. **Emissie(bron)**

Een emissiebron is zowel een activiteit als een inrichting en in het kader van IIOA’s en het Vlarem, representeert het alle verschillende niveaus van een (zowel een schoorsteen als een installatie als een bedrijf op zich). In de UsageNote werd een verwijzing naar IIOA’s van VLAREM toegevoegd.

De attributen (type, ontvanger, locatie, verantwoordelijke, wettelijkkader en identificator) werden toegelicht aan de hand van een voorbeeld op slide 18. We verwijzen dan ook naar deze slide voor meer informatie.

Tijdens de werkgroep werd volgende vraag gesteld:

* Is het datatype van de attributen telkens een vrij in te vullen tekstveld of zitten hier keuzelijsten achter?

*Deze datatypes komen voort uit OSLO Datatypes en er werd voor deze attributen dus niet gewerkt met keuzelijsten buiten bij Emissiebrontype. Voor Emissiebrontype zal nog een keuzelijst voor opgesteld worden. Deze lijst is ook uitbreidbaar.*

*Emissiebrontype kan meerdere classificaties hebben die ook gelinkt zijn aan elkaar. Bijvoorbeeld: de emissiebron is zowel van het type IIOA als van het type fabriek. Deze types kunnen bijvoorbeeld voortkomen uit het VLAREM. In de specificatie zal er aangeduid worden dat dit technisch om een SKOS-Concept codelijst gaat.*

De klasse Emissie werd toegelicht. Hieromtrent kwamen volgende vragen:

* Wat zijn emissietypes?

*Emissietype doelt op de aard van de emissie. De twee belangrijkste emissietypes zijn diffuus en geleid. Deze zullen toegevoegd worden aan de enumeratie.*

* Wat wordt er bedoeld met de matrix van Emissie?

*Het is de omgeving waarin de emissie zich voordoet. Denk hierbij aan lucht, water, grond, grondwater en dergelijke meer. Deze lijst is niet limitatief.*

* + 1. **HydromorfologischIndexType**

[we verwijzen naar de slides voor meer informatie.]

Er werden geen opmerkingen gegeven op de aangepaste enumeratie.

* + 1. **Meetpunt en meetnet**

Zowel de klasse meetpunt als meetnet werd toegevoegd aan het model. Dit om de meetnetten binnen de wereld van de waterkwaliteit (bv. afvalwater) in het model te kunnen brengen. Een opmerking hierbij is dat meetnetten niet uitsluitend uit meetpunten bestaan, maar uit ruimtelijkeBemonsteringsObjecten. Bovendien zou er een exhaustieve lijst gemaakt kunnen worden van de types meetpunten, maar dit valt uit de scope van dit traject.

Als men zou willen weten welk bedrijf gerelateerd is aan het meetpunt, dan kan dit via de klasse ruimtelijkbemonsteringsobject met een een verwijzing naar het bemonsterdobject.

Vanuit de werkgroep kwam de vraag of het mogelijk is om een overzicht te verkrijgen van de IIOA’s en welke emissiepunten deze bevatten. Aan de deelnemer in kwestie werd gevraagd om hierover een issue aan te maken in [github](https://github.com/Informatievlaanderen/OSLOthema-Waterkwaliteit/issues), aangezien het verder uit modelleren van emissiebronnen in deze werkgroep waar observaties centraal staan, te ver zou leiden. **Aggregaties: somparameters en statistiek**

Op basis van een bestaande standaard, namelijk [Uncert](https://www.researchgate.net/publication/228647117_UncertML_an_XML_schema_for_exchanging_uncertainty), werd OSLO Statistiek uitgewerkt om statistieken te kunnen toevoegen aan observaties. Het doel hiervan is om bijvoorbeeld te kunnen zeggen: De gemiddelde concentratie in 2019 van Boor is 50 mg/L, waarbij men expliciet aangeeft in het model dat het om een gemiddelde gaat. Bovendien kan men aan de hand van deze standaard vermelden op welke manier de statistiek is berekend. Hierdoor kan er in de metadata beschreven worden hoe men met uitzonderingen omgaat.

Beide opties die werden gepresenteerd (zie slides) zijn twee gebruiken van hetzelfde, enkel op een andere manier gemodelleerd.

Vanuit de werkgroep kwamen volgende vragen:

* Is het gemiddelde iets dat je publiceert of iets dat je ‘live’ berekend?

*Hier wordt het gemiddelde gepubliceerd.*

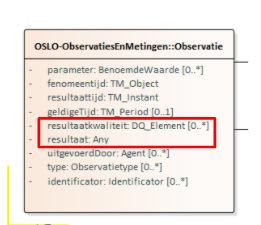
* Kan een chemischAgensConcentratie geen gemiddelde hebben?

*Dit kan wel door double typing. In het geval van een index-observatie (HydromorfologischeIndexObservatie en BiotischeIndexObservatie) weet je dat dit een statistiek is.*

Opmerking: niet het gehele model van [Uncert](https://www.researchgate.net/publication/228647117_UncertML_an_XML_schema_for_exchanging_uncertainty) werd overgenomen omdat dit buiten de scope van dit traject valt.

* + 1. **Kwaliteit**

Er werd een applicatieprofiel OSLO Datakwaliteit toegevoegd op basis van de ISO DataQuality standaard. Dit om meer te kunnen zeggen over de kwaliteit van het resultaat.



* 1. **Toelichtingen van het datamodel**
     1. **Kenmerktype & chemische observaties**

Er werden drie opties gegeven om over een stof te praten:

* SKOS-Concept verwijzing naar stof in een codelijst;
* Rechtstreeks verwijzen naar de stof als instantie van de klasse;
* Het attribuut agens toevoegen aan de klassen ChemischAgensConcentratieObservatie en ChemischAgensVrachtObservatie en het attribuut soort bij BioIndicatorObservatie om zo de klasse te kunnen instantiëren en met behulp van een SKOS-Concept te kunnen verwijzen naar een codelijst. Hier kan men eventueel nog met een SeeAlso verwijzen naar een (internationale) standaard.

Vanuit de werkgroep kwamen volgende opmerkingen:

* Moet de multipliciteit van het kenmerk Agens niet 1 zijn (hierdoor maak je het kenmerk verplicht) aangezien er steeds een Agens geobserveerd zal worden?

*Dit klopt. De multipliciteit van het attribuut Agens zal aangepast worden naar 1.*

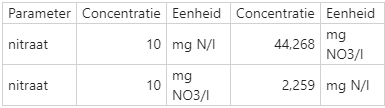
* Dient het attribuut ChemischAgensType bij ChemischAgensVrachtObservatie niet ChemischAgnesKenmerkType te zijn?

*Akkoord. Dit zal zo aangepast worden in het model.*

* Is het mogelijk om in het model bij de eenheid onmiddelijk het chemischAgens mee te nemen zodanig dat deze omrekening niet in een latere fase moet gebeuren?

*Naar modellering toe is het beter om eenheid en chemischAgens van elkaar te scheiden. Dit naar hergebruik en leesbaar toe. Naar rapportering toe is het te begrijpen dat het eenvoudig zou zijn mocht deze omrekening onmiddellijk gedaan worden. In het datamodel van OSLO Waterkwaliteit werd er geopteerd om chemischAgens en eenheid nog steeds van elkaar te scheiden.*

Er werd volgend voorbeeld gedeeld door de werkgroep, waarbij NO3 (Nitraat) uitgedrukt wordt in de eenheid mg N (Stikstof) per l. Hiervoor zal eventueel een [issue](https://github.com/Informatievlaanderen/OSLOthema-Waterkwaliteit/issues) aangemaakt worden op GitHub op welke manier zulke eenheid een plaats heeft in het model.



* + 1. **Beoordeling van waterkwaliteit**

Door de scope van het traject is er besloten om een normoverschrijding te modelleren via twee observaties. De ene observatie is de ‘gewone’ observatie; het is de observatie van het geobserveerdkenmerk. De tweede observatie toetst de ‘gewone’ observatie aan norm waarvan het resultaat waar of onwaar is, namelijk de norm is overschreden of de norm is niet overschreden. In de publieke review periode moet er gezegd worden dat dit niet meer volstaat aan de hand van een issue.   
  
Een alternatief zou zijn om specifiek iets uit te modelleren in het model dat het vergelijken van normen toelaat.

* + 1. **ChemischAgensVracht**

[We verwijzen naar de slides voor meer informatie.]

De relatie tussen een emissie en zijn emissiebron met een chemischAgensVrachtObservatie werd aangetoond.

1. **Volgende stappen**

## Overzicht traject

## Publieke review

De volgende stap in het traject is de publieke review van de standaard. Follow-up communicatie volgt met de te-reviewen standaard, verslag en bijkomende documenten. De looptijd van de publieke review zal nog meegedeeld worden.

De publieke review bestaat uit de volgende onderdelen:

* Publiceren van de standaard met definities en voorbeelden met ingevulde data van de use cases
* Capteren van feedback
* Proof-of-concept

## Feedback tijdens publieke review

Eens de standaard gepubliceerd is en de communicatie van de start van de publieke review is doorgegeven, kan er feedback op de gepubliceerde standaard worden gegeven via GitHub door het aanmaken van issues:

<https://github.com/Informatievlaanderen/OSLOthema-Waterkwaliteit/issues>

## Methode publicatie standaard

Een standaard doorgaat verschillende stadia:

* **Standaard in ontwikkeling**: Het model wordt herzien en de aanvraag wordt ingediend bij de werkgroep datastandaarden om naar publieke review te gaan. Indien aanvaard -> kandidaat standaard in publieke review.
* **Kandidaat standaard:** Aan het einde van de publieke review worden de nodige aanpassingen gemaakt en wordt er een afsluitend webinar georganiseerd.
* **Erkende standaard:** De standaard wordt voorgelegd op de werkgroep datastandaarden, en bij akkoord wordt deze als erkende standaard beschouwd en gepubliceerd op data.vlaanderen.be.

Er wordt zowel een **applicatieprofiel** als een **vocabularium** gepubliceerd. Het vocabularium bevat alle klassen en eigenschappen die gebruikt worden in het applicatieprofiel die nog niet elders in bestaande standaarden worden gebruikt.

**Vervolgtraject Codelijst**

Een van de volgende stappen is een vervolgtraject om de codelijsten uit te werken in het kader van dit traject:

* Chemische stoffen
* Soorten dieren
* Observatieprocedures
* etc.

Hiervoor zal nog communicatie uitgestuurd worden naar de stakeholders die in aanraking komen met de kwaliteit van bodem, water en lucht.