**///** /Thematic workshop 2 OSLO Air & Water/

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

**Datum**: 04/03/2021: 13:00 – 15:30 CET

**Locatie**: Online - Microsoft Teams Meeting

|  |  |
| --- | --- |
| **Belgian national & regional governments** | Michiel De Keyzer - Digitaal Vlaanderen  Kevin Haleydt - Digitaal Vlaanderen Frederik Van den Houdt, Digitaal Vlaanderen Jurgen Meirlaen, Vlaamse Milieumaatschappij  Geert Van Haute, Departement Omgeving  Annelies De Craene, Digitaal Vlaanderen  Greet Devriese, Vlaamse Milieumaatschappij  Frank Lavens, Vlaamse Milieumaatschappij  Geert Thijs, OSLO-team, Digitaal Vlaanderen Raf Buyle, Digitaal Vlaanderen |
| **Local administration Europe** | Benjamin Gärtner, City of Heidelberg, Germany |
| **Research institutions** | Brecht Van de Vyvere, imec-ugent, Belgium  Frank Sleeuwaert, VITO, Belgium  Fernando López, FIWARE Foundation, Germany |
| **Other** | Stijn Van Hoey, Fluves, Belgium  Niels Melotte, De Vlaamse Waterweg, Belgium  Elien Dewitte - VLIZ - Beligum Laurian Van Maldeghem, Flanders Marine Institute (VLIZ),  Belgium  Gert De Tant - ODALA  Laurens Horvath – ALTIS  Olivier Bernard, ALTIS Groupe SA, Switzerland |

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

# Agenda of the workshop

|  |  |  |
| --- | --- | --- |
| **Topic** | **Time (mins)** | **Timestamps** |
| **Start TW2** | 5 mins | 13:05 |
| **Introduction & Who’s in the room** | 10 mins | 13:15 |
| **Recap of Thematic Workshop 1** | 5 mins | 13:20 |
| **Changes since TW1 - Core model** | 15 mins | 13:35 |
| **Changes since TW1 - Air & Water** | 10 mins | 13:45 |
| **Break** | 15 mins | 14:00 |
| **Breakout Session & Debrief** | 30 mins | 14:30 |
| **Specific Models brought up in TW1** | 15 min | 14:45 |
| **Questions for changes to the model** | 30 min | 15:15 |
| **Definitions & Next steps** | 15 min | 15:30 |

1. **Introduction & who’s in the room**

We started the workshop with detailing the purpose of the workshop and the agenda (above). The purpose of the workshop was fourfold, namely:

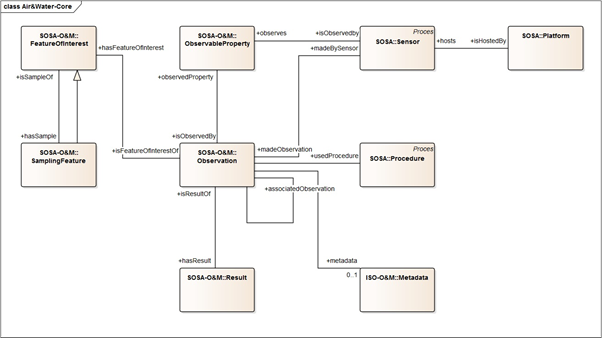
* Giving a recap of the First thematic workshop;
* Going over the changes to core, air & water model since last workshop;
* Addressing specific questions and remarks asked in TW1; and
* Detailing the collaboration for definitions.



1. **Recap of TW1**

Last workshop, we proposed an initial draft model and captured your feedback. The main points that were brought up last workshop were:

* Overall, there was an agreement on the fundamental building blocks (classes) proposed in the model
* With regards to ObservableProperty there were further questions on the inclusion of (micro-) organisms, but also inclusion of pollution from for example a sewer system/chimney
* Additionally, there were specific questions on which units of measurement to be used. Options such as QUDT were proposed
* Questions on the sampling feature and whether there would be a way to document sampling method
* The need to look more in detail into models such as:
  + W3 Cube
  + WISE
  + WaterML Water Quality – WaterML2
  + QUDT



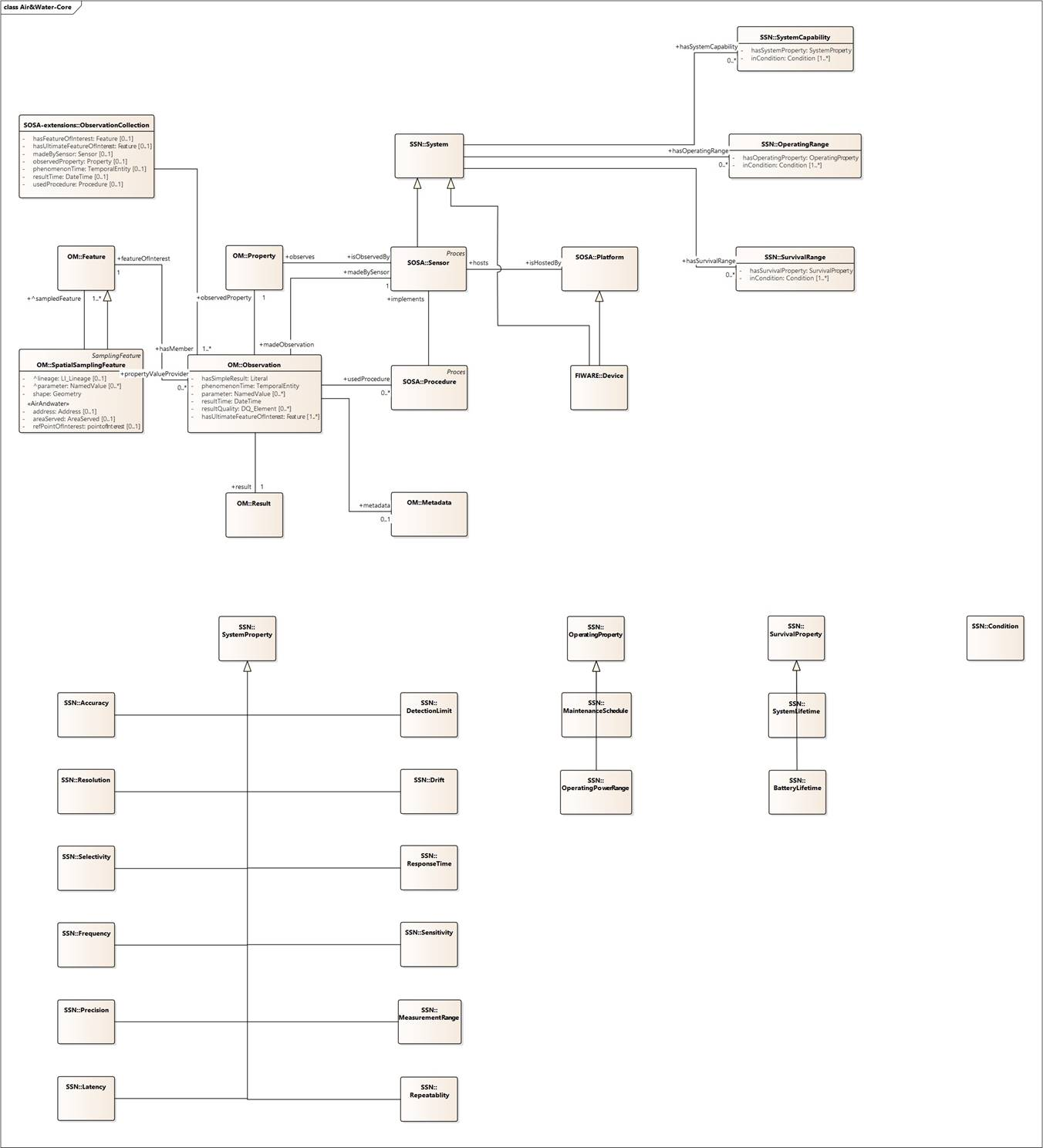
1. **Changes to the core, Air & Water model**

Core Model

In total 4 main changes were made to the core model. These changes were presented in the workshop and agreed upon:

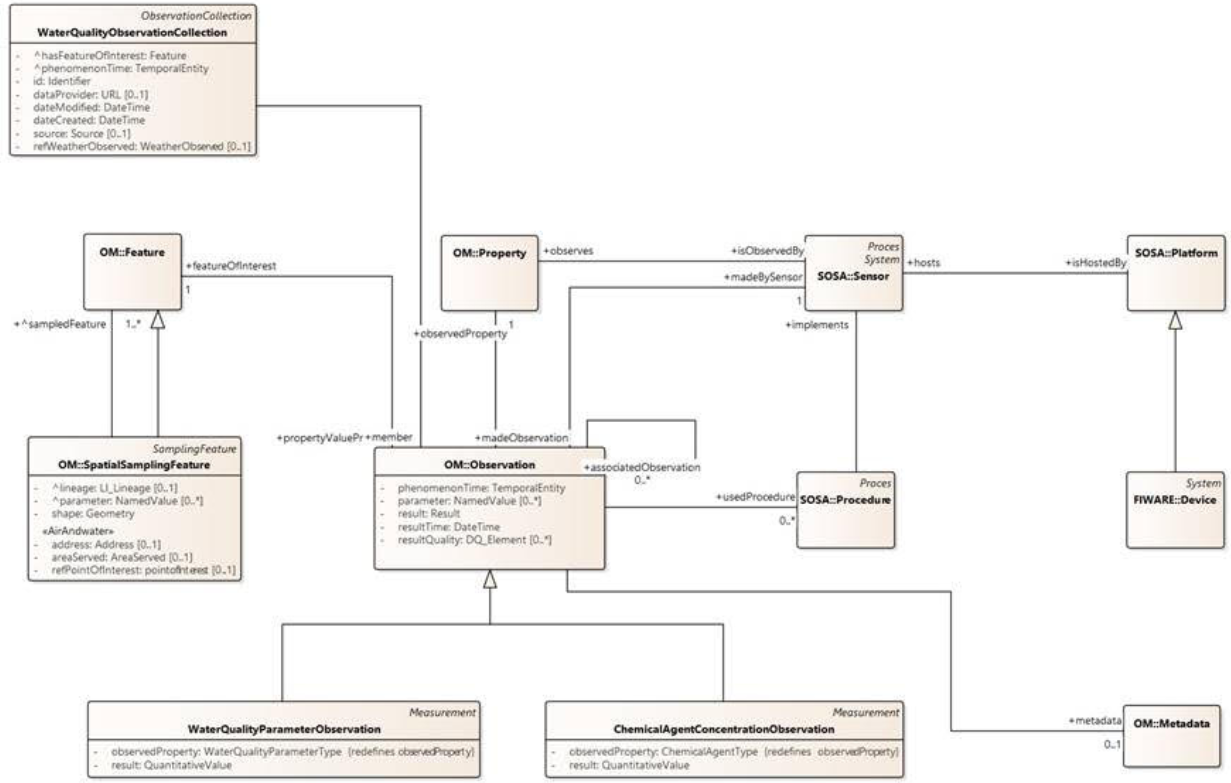
1. ISO O&M to be preferred over SOSA interpretation for the following classes: **Feature of Interest, Observable Property, Sampling Feature, Observation & Results.** The main reason for this choice was that the definitions of ISO O&M were better aligned and had attributes that were ready to be reused. Additionally, a modelling issue was present in the SOSA interpretation of Feature of Interest & Sampling Feature. Ultimately, most other models are building on top of ISO O&M so the logical thing is to do the same.
2. First proposal of attributes for Observation & Spatialsamplingfeature based on ISO O&M
3. Extension of the model on the side of sensor & platform with SSN:System and the accompanying classes such as Systemcapabilities, Operating Range & Survival Range.
4. Addition of SOSA-extensions **ObservationCollection**. This class allows to make aggregations or groups of observations.

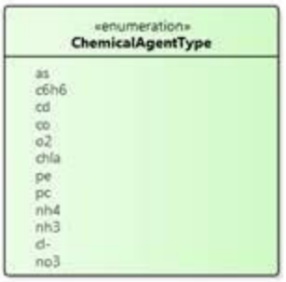
Below you can find the full core model with all the changes listed above.



Water Model

After presenting the core model, a first presentation was made of the water model. This is mainly a copy of the core model, with exception of the Result attribute. This has been replaced with two further specializations of OM Observation, namely:

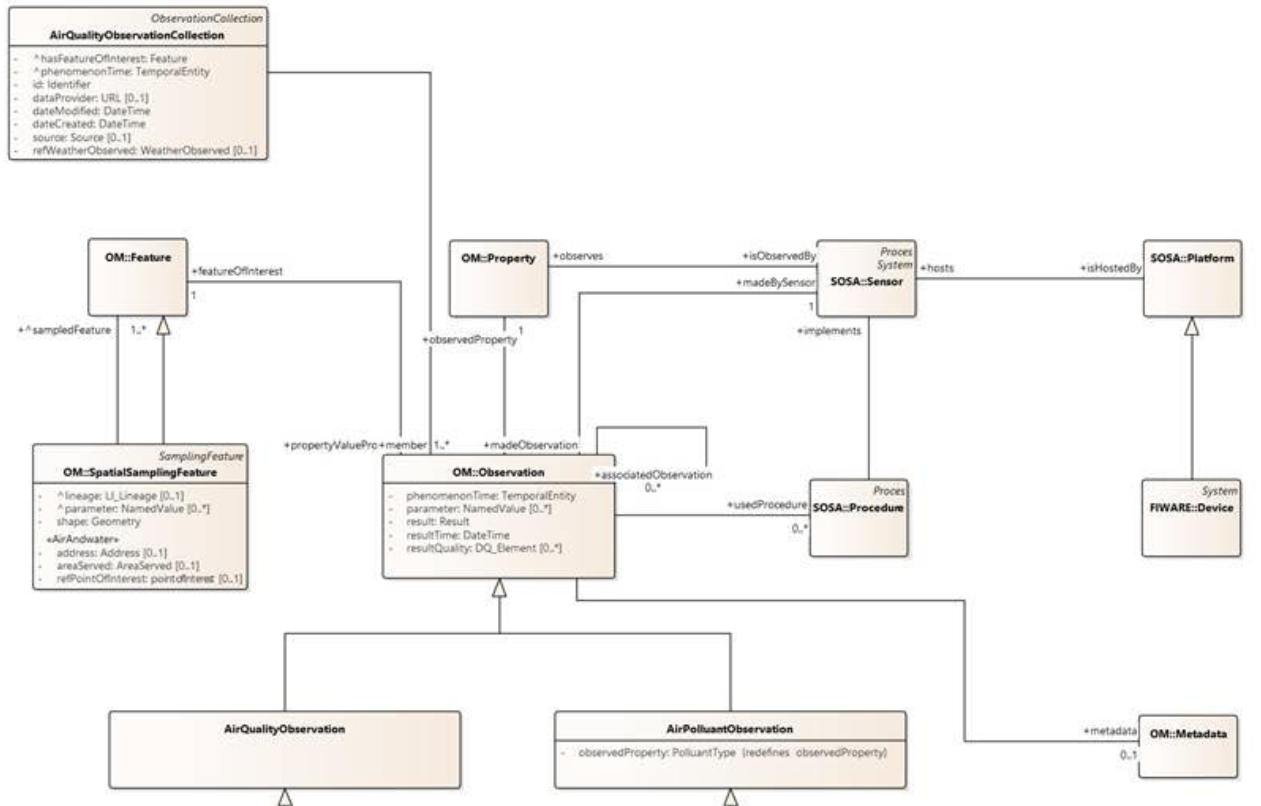
* **Water Quality Parameter Observation:** Types of parameter observations are for example temperature, conductivity, …
* **Chemical Agent Concentration Observation:** Describes the observations of mainly chemical agents such as NH3, CO, …



A more practical object diagram can be found in the presentation or on the dedicated [Github page](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/tree/master/Thematic%20Workshop%202).

Air Model

Similar to the water model, the air model can be seen as a copy of the core model, with the exception of the Result attribute. Here, the results are also modelled as further specializations of the Observation class:

* **AirQualityIndexObservation:** This is a measurement of an air quality index
* **AirQualityLevelObservation:** This is a categorical observation of the air quality level
* **AirPollutantObservation:** This corresponds to an air pollutant value (measurement)
* **AirPollutantLevelObservation:** This corresponds to a categorical representation of the air pollutant level.

A more practical object diagram can be found in the presentation or on the dedicated [Github page](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/tree/master/Thematic%20Workshop%202).

1. **Breakout session**

After the presentation of the different models a breakout session was organized in order to capture input from the group on the two models. You can consult the klaxoon with the input captured [HERE](https://app.klaxoon.com/participate/board/H5KQMV8).

Main issues and feedback capture for the Air Model:

* Typo in the model: Polluant to be changed to Pollutant
* Further subdivision of feature may be needed
* The model should be tested in order to see how we can address the use case of Stack (chimney) emissions to air (from industry). A [github issue](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/issues/1) will be logged on this with a detailed explanation of this use case.

Main issues and feedback captured for the Water Model:

* What about samples taken to the lab for analysis with no direct input from a sensor: This is possible. ISO O&M has a class called “process” which adds analytical processes for lab-related information.
* With respect to occurences, Darwin Core might be relevant and needs to be looked at. A Github issue was already logged for this prior to the workshop. We will investigate this for the next workshop.
* Example of an instance of the class OM:Property where you observe an animal with an external URI. How do you add the property ‘isObservedBy’: The property would be “Occurrence (Yes/No)” where animal is the FeatureOfInterest. A ‘count’ subclass of ISM O&M Observation can be used for this.
* Do calibration parameters of a sensor fall under ‘metadata’ of the observation or are they part of the sensor data?
* Are there links possible with published FeaturesOfInterest such as for example [water networks](http://www.geopunt.be/catalogus/datasetfolder/7eefed2f-5fc9-4d81-8adf-728f9c54aa4a) or [sewer networks](http://www.geopunt.be/catalogus/datasetfolder/6bf3ccbf-9606-4b84-9c1e-53162bf4c737)
* WaterQualityBiologyObservation is missing: This will be addressed in the next workshop.
* Why not generalize the distinction between value and level assumed for the Air model, it can also make sense for Water? There was no ‘levels’ designed in the FIWARE model. In the EU model however, they talk about ‘classes’
* Can we use the ObservationCollection class to build timeseries? This is indeed possible. Another possibility can be to model the results as a timeseries like for example done in ISO O&M (SpatialCoverage) or WaterML 2 through Time-value pairs or Domain Range timeseries.

Other questions and information captured throughout the workshop

We also captured the following discussion points during the workshop:

* Certain capabilities such as latency can not only be linked to a specific system, but also to a platform. How can we split this?
* Look into grouping of sensors for example by cleaning cycle. Can this be covered in ObservationCollection or do we need to alter the model to allow this?
* Modelling of “input” of an observation
* ChemicalAgentConcentration should be seen more broadly. The semantics do not cover for example species or (micro)organisms
* Model observedProperty as a property that could be a coded property
* With regards to unit of measurement, [UN/CEFACT approach](https://unece.org/DAM/cefact/recommendations/rec20/Rec20rev14e-Annex_II-III_2020.xls) can be of use. This will need to be investigated
* Feature of interest: WISE only considers surface water, there are specific types missing such as sewer water, wastewater…
* CAS numbers for chemical substances, however some other measured parameters do not have a CAS number but an EEA number. This is for example the case for measured properties like conductivity, …
  + Need for a layer of abstraction
  + Pubchem is a good published source of substances. Can be useful for observed properties.
* Other links and information share:
  + [Smart Data Models Device](https://github.com/smart-data-models/dataModel.Device/blob/master/Device/examples/example-normalized.jsonld)
  + [Occurences](https://dwc.tdwg.org/terms/)
  + [Wikidata](https://www.wikidata.org/wiki/Q27855)
  + [PubChem for chemical substansces](https://pubchem.ncbi.nlm.nih.gov/compound/Phosphoric-acid)
  + [Webservice](https://www.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pccompound&retmax=100&term=120-83-2) for translating CAS numbers to pubchem IDs

1. **Specific models brought up in TW1**

In TW1, a couple of international models were brought up that may be of interest for the Air & Water models. These were looked into and presented during TW2. More information on these models can be found in the PowerPoint presentation.

1. **Questions for changes to the models**

From our end, we have a few topics we want to bring to the group’s attention and would like to see answered. Based on the input we will then change the model accordingly. A Github issue will be created for each of these questions and you will be able to give feedback on each specific issue. Please take your time to review these and give your input on these matters.

1. [ObservedProperties](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/issues/2)

We will have to foresee standardized lists of observedProperties. The possible options are to:

* Leave the lists to be customizable to whoever exchanges data, as long as the properties are identified via a URI (e.g. SKOS:Concept)
* Suggest standardized lists in the specifications
* Create our own taxonomy or list from what is available

Certain standardized lists have already been brought forward for water ([WISE](http://dd.eionet.europa.eu/dataelements/75873)), biology ([Darwin Core](https://www.gbif.org/dataset/search?type=CHECKLIST)). Are there other standardized lists (mainly for Air).

1. [Further modelling of FeaturesOfInterest](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/issues/3)

There may be a need to further model the Feature of Interest for both Air & Water.

For **water**, [WISE](http://dd.eionet.europa.eu/dataelements/75907) suggests the following subdivision:

* Coastal Water
* Groundwater body
* Lake water body
* Marine waters
* River water body
* Territorial waters
* Transitional water body

The question remains whether this is a full list or not (sewers or water networks). Are there other more complete lists available in other standards? Do we model this or add it as a code list? The main downside for choosing a code list over modelling the different types will be the lack of adding specific attributes to the different Features of interest.

For **Air** the same question arises. Do we need a further extension of the Feature of Interest. And if so, based on which subdivision. Are there any international models that go deeper into this? From the working group, at first glance it seemed like a further extension might be needed to address the use case of stack emissions from chimneys.

1. [SamplingFeatures](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/issues/4)

We suggest adding an extra class that allows tracking ‘**Specimen’**. There are two possibilities:

* The WISE model allows pointing towards physical samples by creating two classes: SpatialSamplingFeatures, Specimen.
* SOSA goes more into details by adding SOSA Sampling (the activity of sampling) and SOSA Sampler (the specific instruments used).

1. [Observations](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/issues/5)

We suggest expanding Observations with other subtypes on top of Measurements. ISO O&M suggests the following subtypes:

* Measurement (value + unit)
* CategoryObservation (classification)
* CountObservation (count)
* TruthObservation (Boolean)
* TemporalObservation (when does the observed characteristics occur)
* GeometryObservation (where does the observation occur)
* ComplexObservation (complex results)
* DiscreteCoverageObservation (timeseries)

1. [ObservationCollections](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/issues/6)

There are multiple ways of depicting variation of data in one or more dimensions:

* W3 Cube was referenced in TW1.
* SOSA extensions Observation collections. This has already been added as part of the core model.

The current implementation of SOSA extensions ObservationCollections allows to make aggregations of data in 1 or more dimensions through the use of Feature of interest/ultimate feature of interest, phenomenon time, observed property, sensor, used procedure. The addition of an identifier to the observation collection will also allow any other arbitrary grouping of observations.

Is this enough to group observations? Do we need to add other aggregations?

1. **Next steps**

* The next workshop is planned on April 1st, 2021 from 13h to 15h30. Don’t forget to subscribe if you haven’t already via the following [**LINK**](https://overheid.vlaanderen.be/informatie-vlaanderen/agenda/thematic-workshop-3-oslo-air-water).
* If you have feedback on the models or other issues that you think should be addressed, please log these as issues on Github ([**https://github.com/Informatievlaanderen/OSLOthema-airAndWater**](https://github.com/Informatievlaanderen/OSLOthema-airAndWater))or send them via email to
  + [kevin.haleydt@vlaanderen.be](mailto:kevin.haleydt@vlaanderen.be)
  + [frederik.van.den.houdt@pwc.com](mailto:frederik.van.den.houdt@pwc.com)
* On our end, we will log the 5 questions addressed in Section 6. QUESTIONS FOR CHANGES TO THE MODELS on Github as well. Please provide your input on these questions and let us know whether you agree with the proposed changes or not.
* You will also find a first list of definitions [HERE](https://github.com/Informatievlaanderen/OSLOthema-airAndWater/tree/master/Thematic%20Workshop%202). Please provide your feedback on these. We will consolidate you feedback and present the results of this on the next workshop.