



# Spatial Data on the Web Use Cases & Requirements

W3C Working Group Note 17 December 2015

**This version:**

<http://www.w3.org/TR/2015/NOTE-sdw-ucr-20151217/>

**Latest published version:**

<http://www.w3.org/TR/sdw-ucr/>

**Latest editor's draft:**

<http://w3c.github.io/sdw/UseCases/SDWUseCasesAndRequirements.html>

**Previous version:**

<http://www.w3.org/TR/2015/NOTE-sdw-ucr-20150723/>

**Editors:**

Frans Knibbe, [Geodan](#)

Alejandro Llaves, [OEG, Universidad Politécnica de Madrid](#)

**OGC Document Number:**

OGC 15-074 R1

**Repository:**

[We are on GitHub](#)

**Changes:**

[Change log](#)

[Diff to previous version](#)

[Commit history](#)

Copyright © 2015 [OGC](#) & [W3C](#)® ([MIT](#), [ERCIM](#), [Keio](#), [Beihang](#)), [W3C liability, trademark](#) and [document use](#) rules apply.

---

## Abstract

This document describes use cases that demand a combination of geospatial and non-geospatial data sources and techniques. It underpins the collaborative work of the [Spatial Data on the Web Working Groups](#) operated by both [W3C](#) and [OGC](#).

## Status of This Document

*This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current [W3C](#) publications and the latest revision of this technical report can be found in the [W3C technical reports index](#) at <http://www.w3.org/TR/>.*

**For OGC** This is a Public Draft of a document prepared by the Spatial Data on the Web Working Group (SDWWG) - a joint [W3C](#)-OGC project (see [charter](#)). The document is prepared following [W3C](#) conventions. The document is released at this time to solicit public comment.

This document was published by the [Spatial Data on the Web Working Group](#) as a Working Group Note. If you wish to make comments regarding this document, please send them to [public-sdw-wg@w3.org](mailto:public-sdw-wg@w3.org) ([subscribe](#), [archives](#)). All comments are welcome.

Publication as a Working Group Note does not imply endorsement by the [W3C](#) Membership. This is a draft document and may be updated, replaced or obsoleted by other documents at any time. It is inappropriate to cite this document as other than work in progress.

This document was produced by a group operating under the [5 February 2004 W3C Patent Policy](#). [W3C](#) maintains a [public list of any patent disclosures](#) made in connection with the deliverables of the group; that page also includes instructions for disclosing a patent. An individual who has actual knowledge of a patent which the individual believes contains [Essential Claim\(s\)](#) must disclose the information in accordance with [section 6 of the W3C Patent Policy](#).

This document is governed by the [1 September 2015 W3C Process Document](#).

## Table of Contents

1. Introduction
2. Deliverables
  - 2.1 Use Cases and Requirements
  - 2.2 Spatial Data on the Web Best Practices
  - 2.3 Time Ontology in OWL
  - 2.4 Semantic Sensor Network Vocabulary
  - 2.5 Coverage in Linked Data
3. Methodology
4. Use Cases
  - 4.1 Meteorological data rescue
  - 4.2 Habitat zone verification for designation of Marine Conservation Zones
  - 4.3 Real-time wildfire monitoring
  - 4.4 Diachronic burnt scar Mapping
  - 4.5 Harvesting of Local Search Content
  - 4.6 Locating a thing
  - 4.7 Publishing geographical data
  - 4.8 Consuming geographical data in a Web application
  - 4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities
  - 4.10 Publishing geospatial reference data
  - 4.11 Integration of governmental and utility data to enable smart grids
  - 4.12 Using spatial data from the Web in [GIS](#) systems during emergency response operations
  - 4.13 Publication of air quality data aggregations
  - 4.14 Publication of transport card validation and recharging data
  - 4.15 Combining spatial RDF data for integrated querying in a triplestore
  - 4.16 Dutch Base Registry
  - 4.17 Publishing Cultural heritage data
  - 4.18 Dissemination of 3D geological data
  - 4.19 Publication of raw subsurface monitoring data
  - 4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches
  - 4.21 Driving to work in the snow
  - 4.22 Intelligent Transportation System
  - 4.23 Optimizing energy consumption, production, sales and purchases in Smart Grids
  - 4.24 Linked Data for tax assessment
  - 4.25 Images, e.g. a time series of a water course
  - 4.26 Droughts in geological complex environments where groundwater is important
  - 4.27 Soil data applications
  - 4.28 Bushfire response coordination centre
  - 4.29 Observations on geological samples
  - 4.30 Spatial sampling
  - 4.31 Select hierarchical geographical regions for use in data analysis or visualisation

- 4.32 Satellite data processing
- 4.33 Marine observations - eMII
- 4.34 Marine observations - data providers
- 4.35 Marine observations - data consumers
- 4.36 Building information management and data sharing
- 4.37 Landsat data services
- 4.38 Metadata and search granularity
- 4.39 Crowdsourced earthquake observation information
- 4.40 TCGA / microscopy imaging
- 4.41 Crop yield estimation using multiple satellites
- 4.42 Geospatial extensions to domain-independent metadata schemas
- 4.43 Improving discovery of spatial data on the Web
- 4.44 INSPIRE compliance using Web standards
- 4.45 Event-like geographic features
- 4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model
- 4.47 Incorporating geospatial data (e.g. geo-referenced geometry) into interactive 3D graphics on the Web
- 4.48 Smart Cities
- 4.49 Provenance of climate data
- 4.50 Representing geospatial data in RDF
- 5. Accepted requirements
  - 5.1 Bounding box and centroid
  - 5.2 Compatibility with existing practices
  - 5.3 Compressible
  - 5.4 Coverage temporal extent
  - 5.5 Crawlability
  - 5.6 CRS definition
  - 5.7 Date, time and duration
  - 5.8 Default CRS
  - 5.9 Different time models
  - 5.10 Discoverability
  - 5.11 Dynamic sensor data
  - 5.12 Encoding for vector geometry
  - 5.13 Ex-situ sampling
  - 5.14 Georectification
  - 5.15 Georeferenced spatial data
  - 5.16 Humans as sensors
  - 5.17 Independence on reference systems
  - 5.18 Lightweight API
  - 5.19 Linkability
  - 5.20 Machine to machine
  - 5.21 Mobile sensors
  - 5.22 4D model of space-time
  - 5.23 Model reuse
  - 5.24 Moving features
  - 5.25 Multilingual support
  - 5.26 Multiple types of coverage
  - 5.27 Nominal observations
  - 5.28 Nominal temporal references
  - 5.29 Non-geographic reference system
  - 5.30 Observation aggregations
  - 5.31 Observed property in coverage
  - 5.32 Provenance
  - 5.33 Quality metadata
  - 5.34 Reference data chunks
  - 5.35 Reference external vocabularies
  - 5.36 Sampling topology
  - 5.37 Sensor metadata

- 5.38 Sensing procedure
- 5.39 Space-time multi-scale
- 5.40 Spatial metadata
- 5.41 Spatial relationships
- 5.42 Spatial operators
- 5.43 Spatial vagueness
- 5.44 SSN-like representation
- 5.45 Streamable data
- 5.46 Support for 3D
- 5.47 Time dependencies in CRS definitions
- 5.48 Temporal reference system
- 5.49 Temporal vagueness
- 5.50 Support for tiling
- 5.51 Time series
- 5.52 Uncertainty in observations
- 5.53 Update datatypes in OWL Time
- 5.54 Use in computational models
- 5.55 Validation
- 5.56 Valid time
- 5.57 Virtual observations
- 6. Accepted requirements by deliverable
  - 6.1 Requirements for deliverable Spatial Data on the Web Best Practices:
  - 6.2 Requirements for deliverable Time Ontology in OWL:
  - 6.3 Requirements for deliverable Semantic Sensor Network Vocabulary:
  - 6.4 Requirements for deliverable Coverage in Linked Data:
- 7. Deferred requirements
  - 7.1 Definition of SSN profiles and checking compliance against the model
- 8. Acknowledgements
- 9. Appendix A: History of changes
  - 9.1 Changes between the First Public Working Draft and Second Public Working Draft

## 1. Introduction

The mission of the [Spatial Data on the Web Working Group](#), as described in its [charter](#), is to clarify and to formalize standards on spatial data on the Web. In particular:

1. to determine how spatial information can best be integrated with other data on the Web;
2. to determine how machines and people can discover that different facts in different datasets relate to the same place, especially when 'place' is expressed in different ways and at different levels of granularity;
3. to identify and assess existing methods and tools and then create a set of best practices for their use;
4. where desirable, to complete the standardization of informal technologies already in widespread use.

This document describes the results of the first steps of working towards these goals. Members of the working group and other stakeholders have come up with a number of [use cases](#) that describe how spatial data on the Web could work. From these use cases, a number of [requirements](#) for further work are derived. In this document, use cases, requirements and their relationships are described. Requirements and use cases are also related to the [deliverables](#) of the working group.

The requirements described in this document will be the basis for development of the other four deliverables of the Working Group.

## 2. Deliverables

The deliverables of this Working Group are described in the [Working Group Charter](#). For convenience those deliverables are replicated in this chapter. The charter remains the authoritative source of the definition of deliverables.

### 2.1 Use Cases and Requirements

A document setting out the range of problems that the working groups are trying to solve (this document).

### 2.2 Spatial Data on the Web Best Practices

This will include:

- an agreed spatial ontology conformant to the ISO 19107 abstract model and based on existing available ontologies such as GeoSPARQL, NeoGeo and the ISA Core Location vocabulary;
- advice on use of URIs as identifiers in GI systems;
- advice on providing different levels of metadata for different usage scenarios (from broad sweep metadata to metadata about individual coordinates in a polygon);
- develop advice on, or possibly define, RESTful APIs to return data in a variety of formats including those defined elsewhere, such as GeoJSON, GeoJSON-LD and TopoJSON.

### 2.3 Time Ontology in OWL

The WG will work with the authors of the existing Time Ontology in OWL to complete the development of this widely used ontology through to Recommendation status. Further requirements already identified in the geospatial community will be taken into account.

### 2.4 Semantic Sensor Network Vocabulary

The WG will work with the members of the former Semantic Sensor Network Incubator Group to develop its ontology into a formal Recommendation, noting the work to split the ontology into smaller sections to offer simplified access.

### 2.5 Coverage in Linked Data

The WG will develop a formal Recommendation for expressing discrete coverage data conformant to the ISO 19123 abstract model. Existing standard and de facto ontologies will be examined for applicability; these will include the RDF Data Cube. The Recommendation will include provision for describing the subset of coverages that are simple timeseries datasets - where a time-varying property is measured at a fixed location. OGC's WaterML 2 Part 1 - Timeseries will be used as an initial basis.

Given that coverage data can often be extremely large in size, publication of the individual data points as Linked Data may not always be appropriate. The Recommendation will include provision for describing an entire coverage dataset and subsets thereof published in more compact formats using Linked Data. For example where a third party wishes to annotate a subset of a large coverage dataset or a data provider wishes to publish a large coverage dataset in smaller subsets to support convenient reuse.

NOTE

The OGC is currently working on refinements of ISO 19123 (in particular, the OGC Coverage Implementation Schema 1.1), which could result in specifications that allow a higher level of interoperability of implementations. The Working Group will also consider these forthcoming standards.

### 3. Methodology

In order to find out the requirements for the deliverables of the Working Group, use cases were collected. For the purpose of the Working Group, a use case is a story that describes challenges with respect to spatial data on the Web for existing or envisaged information systems. It does not need to adhere to certain standardised format. Use cases are primarily used as a source of requirements, but a use case could be revisited near the time the work of the Working Group will reach completion, to demonstrate that it is now possible to make the use case work.

The Working Group has derived requirements from the collected use cases. A requirement is something that needs to be achieved by one or more deliverables and is phrased as a specification of functionality. Requirements can lead to one or more tests that can prove whether the requirement is met.

Care was taken to only derive requirements that are considered to in scope for the further work of the Working Group. The scope of the Working Group is determined by the [charter](#). To help keeping the requirements in scope, the following questions were applied:

1. Is the requirement specifically about spatial data on the Web?
2. Is the use case including data published, reused, and accessible via Web technologies?
3. Has a use case a description that can lead to a testable requirement?

### 4. Use Cases

Use cases that describe current problems or future opportunities for spatial data on the Web have been gathered as a first activity of the Working Group. They were mainly contributed by members of Working Group, but there were also contributions from other interested parties. In this chapter these use cases are listed and identified. Each use case is related to one or more Working Group [deliverables](#) and to one or more [requirements](#) for future deliverables.

#### 4.1 Meteorological data rescue

**Contributed by:** Chris Little, based on scenarios used for the WMO infrastructure requirements.

**? Full use case description (click to expand):**

This is really one of several future, but realistic, meteorological scenarios to aim at.

National Hydro-Meteorological Services around the world are coordinated via the WMO (World Meteorological Organization), part of the United Nations system. WMO has the same status as ISO, and its standards and regulatory materials applies to all its 193 national meteorological services and are available in the six working languages ( عربي | 中文 | Fr | Ru | Es | En). WMO has embarked on a long-term (think a decade or so) program to update the global meteorological operational infrastructure. This is known as the WIS (WMO Information System). The global infrastructure also has aviation, oceanographic, seismic and other users. The WIS includes a global, federated, synchronized, geospatial catalog, envisaged to encompass all hydro-meteorological data and

services. Currently several nodes are operational, cataloging mainly routinely exchanged observations and forecasts.

Envisage an environmental scientist in Cambodia, researching the impact of deforestation in Vietnam as part of investigating the regional impacts of climate change. She submits her search keywords, in Cambodian, and receives responses indicating there is some data from the 1950s, printed in a 1960 pamphlet, in the Bibliothèque Nationale, a library in Paris, France, in French. She receives an abstract of some form that enables her to decide that the data are worth accessing, and initiates a request for a digital copy to be sent.

She receives the pamphlet as a scanned image of each page, and she decides that the quantitative information in the paper is useful, so she arranges transcription of the tabular numerical data and their summary values into a digital form and publishes the dataset, with a persistent identifier, and links it to a detailed coverage extent, the original paper source, the scanned pages and her paper when it is published. She also incorporates scanned charts and graphs from the original pamphlet into her paper. Her organization creates a catalog record for her research paper dataset and publishes it in the WIS global catalog, which makes it also visible to the GEO System of Systems broker portal.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.40 Spatial metadata](#), [5.4 Coverage temporal extent](#), [5.6 CRS definition](#), [5.7 Date, time and duration](#), [5.9 Different time models](#), [5.10 Discoverability](#), [5.15 Georeferenced spatial data](#), [5.19 Linkability](#), [5.25 Multilingual support](#), [5.28 Nominal temporal references](#), [5.31 Observed property in coverage](#), [5.32 Provenance](#), [5.33 Quality metadata](#), [5.34 Reference data chunks](#), [5.35 Reference external vocabularies](#), [5.38 Sensing procedure](#), [5.37 Sensor metadata](#), [5.39 Space-time multi-scale](#), [5.43 Spatial vagueness](#), [5.48 Temporal reference system](#), [5.52 Uncertainty in observations](#), [5.5 Crawlability](#)

## 4.2 Habitat zone verification for designation of Marine Conservation Zones

**Contributed by:** Jeremy Tandy

**? Full use case description (click to expand):**

The [Marine and Coastal Access Act 2009](#) allows for the creation of a type of Marine Protected Area (MPA), called a Marine Conservation Zone (MCZ). MCZs protect a range of nationally important marine wildlife, habitats, geology and geomorphology and can be designated anywhere in English and Welsh inshore and UK offshore waters.

The designation of a MCZ is dependent on a detailed analysis of the marine environment which results in the definition of geometric areas where a given habitat type is deemed to occur and is published as a habitat map.

Being a policy statement, it is important to be able to express the provenance of information that was used to compile the habitat map. Moreover, because the marine environment is always changing, it is important to express the time at which this information was collected.

The information includes:

- acoustic survey
- video (from a video camera towed behind a survey boat)
- biota observations (based on what is observed in the video and from physical collection)
- particle size (sand/mud)
- water column data
- seabed character map: discrete seabed features and backscatter information (from sonar) to determine bottom type



These information types are varied in type and size. In particular, the acoustic survey (e.g. side-scan sonar) is difficult to manage as these survey results can be many gigabytes in size and cover large areas. A way is needed to refer to just a small part of these coverage data sets that are relevant to a particular habitat zone analysis.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.15 Georeferenced spatial data](#), [5.26 Multiple types of coverage](#), [5.32 Provenance](#), [5.34 Reference data chunks](#), [5.37 Sensor metadata](#), [5.6 CRS definition](#), [5.35 Reference external vocabularies](#), [5.21 Mobile sensors](#), [5.19 Linkability](#), [5.27 Nominal observations](#), [5.16 Humans as sensors](#), [5.46 Support for 3D](#), [5.13 Ex-situ sampling](#), [5.38 Sensing procedure](#), [5.41 Spatial relationships](#)

### 4.3 Real-time wildfire monitoring

**Contributed by:** Manolis Koubarakis

#### ? Full use case description (click to expand):

This use case is about the wildfire monitoring service of the National Observatory of Athens (NOA) as studied in the [TELEIOS](#) project. The wildfire monitoring service is based on the use of satellite images originating from the SEVIRI (Spinning Enhanced Visible and Infrared Imager) sensor on top of the Meteosat Second Generation satellites MSG-1 and MSG-2. Since 2007, NOA operates an MSG/SEVIRI acquisition station, and has been systematically archiving raw satellite images on a 5 and 15 minutes basis, the respective temporal resolutions of MSG-1 and MSG-2.

The service active in NOA before TELEIOS can be summarized as follows:

1. The ground-based receiving antenna collects all spectral bands from MSG-1 and MSG-2 every 5 and 15 minutes respectively.
2. The raw datasets are decoded and temporarily stored in the METEOSAT Ground Station as wavelet compressed images.
3. A Python program manages the data stream in real-time by offering the following functionality:
  1. Extract and store the raw file metadata in an SQLite database. This metadata describes the type of sensor, the acquisition time, the spectral bands captured, and other related parameters. Such a step is required as one image comprises multiple raw files, which might arrive out-of-order.
  2. Filter the raw data files, disregarding non-applicable data for the fire monitoring scenario, and dispatch them to a dedicated disk array for permanent storage.
  3. Trigger the processing chain by transferring the appropriate spectral bands via FTP to a dedicated machine and initiating the following steps: (i) cropping the image to keep only the area of interest, (ii) georeferencing to the geodetic reference system used in Greece (HGRS 87), (iii) classifying the image pixels as "fire" or "non-fire" using appropriate algorithms, and finally (iv) exporting the final product to raster and vector formats (ESRI shapefiles).
  4. Dispatch the derived products to the disk array and additionally store them in a PostGIS database system.

It would be interesting for NOA to see how to use the standards developed by this working group to achieve the following:

- Improve the thematic accuracy of generated products.
- Generate thematic maps by combining the generated products with other kinds of data such as: Corine Land Cover, the Administrative Geography of Greece, OpenStreetMap data and Geonames.



This use case is further discussed in [Real-Time Wildfire Monitoring Using Scientific Database and Linked Data Technologies](#). Some of the data used in the operational service is [available separately](#).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.6 CRS definition](#), [5.14 Georectification](#), [5.19 Linkability](#), [5.26 Multiple types of coverage](#), [5.32 Provenance](#), [5.37 Sensor metadata](#), [5.11 Dynamic sensor data](#), [5.44 SSN-like representation](#)

#### 4.4 Diachronic burnt scar Mapping

**Contributed by:** Manolis Koubarakis

##### ? Full use case description (click to expand):

This use case was studied by the National Observatory of Athens (NOA) in the [TELEIOS](#) project. The burnt scar mapping service is dedicated to the accurate mapping of burnt areas in Greece after the end of the summer fire season, using Landsat 5 TM satellite images. The processing chain of this service is divided into three stages, each one containing a series of modules.

The pre-processing stage is dedicated to (i) identification of appropriate data, downloading and archiving, (ii) georeferencing of the received satellite images, and (iii) cloud masking process to exclude pixels “contaminated” by clouds from the subsequent processing steps.

The core processing stage comprises (i) a classification algorithm which identifies burnt and non-burnt sets of pixels, (ii) a noise removal process that is necessary to eliminate isolated pixels that have been classified wrongfully as burnt, and (ii) converting the raster intermediate product to vector format.

Finally, the post-processing stage consists of (i) a visual refinement step to ensure product thematic accuracy and consistency, (ii) attribute enrichment of the product by overlaying the polygons with geoinformation layers and finally (iii) generation of thematic maps. It would be interesting for NOA to see where the standards to be developed in this working group could be used.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.6 CRS definition](#), [5.14 Georectification](#), [5.19 Linkability](#), [5.26 Multiple types of coverage](#), [5.32 Provenance](#), [5.37 Sensor metadata](#), [5.44 SSN-like representation](#)

#### 4.5 Harvesting of Local Search Content

**Contributed by:** Ed Parsons

##### ? Full use case description (click to expand):

This is a rather generic and broad use case, relevant to Google but clearly also relevant to anyone interested in machine processing of HTML referring to about locations and activities that take place at those locations. Local search providers spend much time and effort creating databases of local facilities, businesses and events.

Much of this information comes from Web pages published on the public Web, but in an unstructured form. Previous attempts at harvesting this information automatically have met with only limited success. Current alternative approaches involve business owners manually adding structured data to dedicated portals. This approach, although clearly an improvement, does not really scale and there are clearly issues in terms of data sharing and freshness.

The information of interest includes:

- the facility's address;
- the type of business/activity;
- opening hours;
- date, time and duration of events;
- telephone, e-mail and Web site details.

Complexities to this include multiple address standards, the differences between qualitative representations of place, and precise spatial co-ordinates, definitions of activities etc.

Ultimately these Web pages should become the canonical source of local data used by all Web users and services.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#),

**Related requirements:** [5.7 Date, time and duration](#), [5.5 Crawlability](#), [5.10 Discoverability](#)

## 4.6 Locating a thing

**Contributed by:** Ed Parsons

**? Full use case description (click to expand):**

With the increasing availability of small, mobile location aware devices the requirement to identify a location human terms is becoming more important. While the determination of sensor in space to a high level of precision is a largely solved problem we are less able to express the location in terms meaningful to humans. The fact that the Bluetooth-LE tracker attached to my bag is at 51.4256853,-0.3317991,4.234500 is much less useful than the description, "Under your bed at home". At others times the location descriptions "24 Bridgeman Road, Teddington, TW11 8AH, UK" might be equally valid, as might "Teddington", "South West London", "England", "UK", "Inside", "Where you left it Yesterday", "Upstairs", "45 minutes from here" or "150 meters from the Post Office".

A better understanding of how we describe places in human terms, the hierarchical nature of places and the fuzzy nature of many geographical entities will be needed to make the "Internet of Things" manageable. A new scale of geospatial analysis may be required using a reference frame based on the locations of individuals rather than a global spherical co-ordinate, allowing a location of your keys and their attached bluetooth tag to be described as "in the kitchen".

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.17 Independence on reference systems](#), [5.47 Time dependencies in CRS definitions](#), [5.20 Machine to machine](#), [5.41 Spatial relationships](#)

## 4.7 Publishing geographical data

**Contributed by:** Frans Knibbe

**? Full use case description (click to expand):**

This use case is for representing the perspective of a party that is interested in publishing data on the Web and wants to do it right with respect to the geographical component of the data. The point of this use case is that it would be good to remove barriers that stand in the way of more spatial data becoming available on the Web.

A data publisher could have the following questions:

1. How should I publish vector data? What is the best encoding to use?
2. How should I publish raster data?
3. How do I make the CRS known?
4. How do I make the spatial resolution/level of detail/accuracy known?
5. Which data publishing software has good support of geographical data types and geographical functions?
6. Which data publishing software has good performance when it comes to spatial operations on data?

From the last two questions it follows that the WG could also be involved in enabling conformance testing and stimulating development of benchmarks for software.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.46 Support for 3D](#), [5.1 Bounding box and centroid](#), [5.2 Compatibility with existing practices](#), [5.41 Spatial relationships](#)

## 4.8 Consuming geographical data in a Web application

**Contributed by:** Frans Knibbe

### ? Full use case description (click to expand):

This use case is somewhat complementary to use case [Publishing Geographical Data](#). It takes the consumer perspective, specifically that of a developer of a Web application that should visualize data and allow some kind of user interaction. The hypothetical Web application has little or no prior knowledge about the data it will encounter on the Web, but should be able to do something meaningful with any spatial data that are encountered, like drawing data on a map or rendering the data in a 3D cityscape.

The point of this use case is that in order for spatial data on the Web to be successful, supply and demand must be balanced to create a positive feedback loop. High quality data must be available in high quantities but those data must also be highly usable for experts as well as non-experts.

A Web application developer could have the following questions:

1. How do I find geographical data on the Web?
2. How can I tell what kind of spatial data I will get? Raster or vector? 2D or 3D?
- 3.
4. Which encoding of vector data can I expect?
5. Which encoding of raster data can I expect?
6. Can I get the data with coordinates that match the coordinate system of my map?
7. What is the spatial extent of this dataset/collection of resources/resource?
8. How can I filter data to get the most appropriate spatial representation of a resource/collection of resources?
9. How can I use spatial data on the Web without having to take a four year academic course first?
10. Which spatial operations does this SPARQL endpoint support?
11. Can I use spatial operations in federated queries?
12. How can I ensure responsiveness of my application (low wait times when accessing data)?

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.1 Bounding box and centroid](#), [5.5 Crawability](#), [5.50 Support for tiling](#), [5.17 Independence on reference systems](#), [5.3 Compressible](#)

## 4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities

**Contributed by:** Frans Knibbe, Karl Grossner

### ? Full use case description (click to expand):

Note this use case shares characteristics with [Publishing Cultural Heritage Data](#).

A research endeavor that has just started tries to stimulate researchers in various fields of the humanities to make research data available in such a way that the data are and remain usable by other researchers, and that the data may be used for purposes other than those envisaged by the original researcher. The emphasis lies on spatiotemporal data because they are nice to visualize (a map with a time slider) and because it is thought that it would be interesting to try to discover patterns in time and/or space in interlinked distributed data sets.

This project has the following aspects that seem relevant to this Working Group:

1. Technologies must be easy to implement for people that generally do not have a high affinity with IT. This goes for data publishing as well as data consumption.
2. References to time and space are often inexact or have shifting frames of reference, so simple encodings like basic geo or ISO 8601 do not suffice.
3. References to time and space do need to be as exact as possible, to enable automatic discovery of spatiotemporal patterns.
4. Datasets do not just need to be published, they need to be easily discovered too, using spatial and/or temporal filters.

Adding examples below relevant to items 2 and 3 above, from one existing scholarly Web application case, which may contribute to a more general (i.e. not necessarily historical) requirement for representing several types of uncertainty: imprecision, probability, confidence. Standards for gazetteers -particularly historical (temporal) ones- are non-existent, although several projects with potentially global reach are underway. It will be helpful to have this Working Group in dialog with developers for such projects as [Pelagios](#), [Library of Congress](#), [Pleiades](#), and Past Place (cf. [Humphrey Southall](#)).

### Spatial

- A set of life path data were developed for a kinship network of 30,000 individual Britons linked by birth and marriage. Spatial data for the locations of life events has several levels of granularity, from street address (10 Downing Street) to country (China). How can spatial containment relationships for places be expressed so that spatial-temporal contemporaneity be calculated?
- References to places in historical works are often limited to toponyms (i.e. absent geometry or precise spatial relations), and qualified by such terms as "near," and "north of." How can these be indexed spatially so as to be discoverable?

### Temporal

- As with spatial data, historical sources contain temporal references at varying granularity. A single data set may contain expressions for exact dates, months, or years -- or ranges containing a mix of any of those.
- Temporal references are frequently inexact, or relational with variable precision. The above referenced data set has a mix, including "around March 1832," "before 1750," "after WW II."

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.28 Nominal temporal references](#), [5.39 Space-time multi-scale](#), [5.49 Temporal vagueness](#), [5.47 Time dependencies in CRS definitions](#), [5.17 Independence on reference systems](#), [5.41 Spatial relationships](#)

## 4.10 Publishing geospatial reference data

**Contributed by:** Clemens Portele

? **Full use case description (click to expand):**

*This use is based on the [European Location Framework \(ELF\)](#).*

Mapping and cadastral authorities maintain datasets that provide geospatial reference data. Reference data is data that a user/developer uses to provide location for her own data (by linking to it), by providing context information about a location (overlying his data over a background map), etc.

A key part of this is persistent identifiers for the published data to allow linking to the reference data. Let's assume that http URIs following the [Cool URI note](#) are used as identifiers.

In ELF — and [INSPIRE](#) — reference data is typically published using a Web service by the national authority. In ELF this is an OGC Web Feature Service. To provide access to the different datasets via a single entry point, all the national services are made available via a proxy Web service that also handles authentication etc. In addition, it is foreseen to publish the reference data in other commonly used Web-based platforms for geospatial data to simplify the use of the data - developers and users can use the tools and APIs they are familiar with.

As a result, the same administrative unit (to pick an example) is basically available via multiple (document) URIs: via the national Web service, the ELF proxy Web service and Web services of the other platforms. Different services will support different representations (GML, JSON, etc). The Web services may not be accessible by everyone and different users will have access to different document URIs.

Which real-world object and document URIs for the administrative unit should be maintained and what does a GET return in order:

- to support linking other data to the administrative unit;
- to deliver the document and representation that the user expects when dereferencing the link?

A related challenge is that today such links are often implicit. For example, a post code or a statistical unit code is a property in the other data, but the link is not explicit like an HTTP URI. What is a good practice to make use of such implicit links? Should they be converted to HTTP URIs to be explicit or are there better ways (e.g. additional context that provide information about the semantics and a pattern how to construct dereferencable URIs)?

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.55 Validation](#), [5.19 Linkability](#), [5.42 Spatial operators](#), [5.15 Georeferenced spatial data](#)

## 4.11 Integration of governmental and utility data to enable smart grids

**Contributed by:** Frans Knibbe

? **Full use case description (click to expand):**

The research project [CERISE-SG](#) aims to integrate data from different domains: government, energy utilities and geography, in order to enable establishment of smart energy grids.

The project has recognized Linked Data as an appropriate concept for integration of data from separate semantic domains. One approach of achieving cross-domain interoperability is to switch from domain-specific semantics to common semantics. For example, the concept of an address has its own definitions in governmental standards and utility standards. Using a common definition improves interoperability.

An example of a domain model that is an international standard in electric utilities is the [Common Information Model \(CIM\)](#). Its data model provides definitions for an important entity: the electricity meter. These meters provide useful data on consumption and production of energy. If it is possible to view these devices as sensors, it could be possible to move from domain specific semantics (CIM) to common semantics (SSN), and to have ready linkage to geographical semantics (location and network topology). What is required in this case is a low-threshold way of using sensor semantics, because people involved in integration of data from multiple domains should not be burdened with having to grasp the full complexity of each domain.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.37 Sensor metadata](#), [5.19 Linkability](#), [5.10 Discoverability](#), [5.40 Spatial metadata](#)

## 4.12 Using spatial data from the Web in GIS systems during emergency response operations

**Contributed by:** Bart van Leeuwen

### ? Full use case description (click to expand):

During emergency response operations in the Netherlands the only spatial data available to emergency response services is the predefined data in their GIS warehouses. Because incidents and accidents handled by emergency response services are by nature unpredictable it is impossible to determine beforehand which data are needed. Ad-hoc data on the Web with a spatial component available are not easily used right now, neither can the available data about a incident be easily shared on the Web.

A related issue is that the relation between the data in the GIS warehouses is only spatial, while most of the time there is an [administrative relation](#) as well.

Under the umbrella of various projects a first attempt has been made to at least share definitions of the terminology used by various emergency response services, both national and cross border. This resulted in the start of a project called the [Firebrary](#), now the terminology and definitions are available on the Web as linked data as SKOS. Still linking from the spatial data to these definitions and vice versa is not standardized.

What seems to be missing is a way to define a relationship between GIS data (in e.g. WFS services) and data on the Web and vice versa, express that a certain linked data resource is also available through a GIS service (rdfs:seeAlso is considered by many to be too limited) or describe the spatial component in the linked data directly.

Being able to plot and exchange data about active incidents through the Web and visualize them in GIS tools with open standards would be a huge leap forward for emergency response services.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)



**Related requirements:** [5.2 Compatibility with existing practices](#), [5.10 Discoverability](#), [5.40 Spatial metadata](#), [5.19 Linkability](#)

#### 4.13 Publication of air quality data aggregations

**Contributed by:** Alejandro Llaves, Miguel Angel García-Delgado (OEG-UPM), Rubén Notivol, Javier Celma (Ayuntamiento de Zaragoza)

? **Full use case description (click to expand):**

**What:** The local authorities of Zaragoza (Spain) want to publish the air quality data of the city. Each observation station has a spatial location described with an address. The dataset contains hourly observations and daily aggregations of different gases, e.g. SO<sub>2</sub>, NO<sub>2</sub>, O<sub>3</sub>, CO, etc.

**How:** We use the [Location Core vocabulary](#) to model the address, e.g. :station locn:address "C/ Gran Vía (Paraninfo)"^^xsd:string. We use xsd:dateTime to represent hourly observations, e.g. :obs ssn:observationResultTime "2003-03-08T11:00:00Z"^^xsd:dateTime.

Open challenges: The combination of hourly observations and daily aggregations in the same dataset may cause confusion because the granularity of the observation is not explicit. For daily aggregations, we suggest using time:Interval from the Time Ontology. To make the temporal modeling more homogeneous, time:Instant could be used for the hourly observations. [\[LINK TO DATA SAMPLE TO BE ADDED\]](#)

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.3 Time Ontology in OWL](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.23 Model reuse](#), [5.30 Observation aggregations](#), [5.35 Reference external vocabularies](#)

#### 4.14 Publication of transport card validation and recharging data

**Contributed by:** Alejandro Llaves (OEG-UPM)

? **Full use case description (click to expand):**

**What:** The Regional Transport Consortium of Madrid (CRTM) wants to make available data about transport card validations and transport card recharging. In the case of transport card validations, the NFC sensors are located on buses, and at the entrance and (some) exit points of metro stations. The observation value of a validation includes data related to the transport card, such as the card identifier and the user profile. The sensors for transport card recharging are ATMs and ticket selling points distributed around Madrid. The observation value of a recharging includes the card identifier and the type of recharging.

**How:** To model transport card validations, we consider two observed properties: user entry (EntradaUsuario) and user exit (SalidaUsuario). Validation sensors at metro stations have a fixed location and a unique identifier, e.g. 02\_L12\_P2. A bus validation sensor is moving continuously, so for the sake of pragmatism, there is a unique sensor identifier for each bus stop in every line, e.g. 03\_L20\_P837. Those identifiers point to an address and geographic coordinates. The observed property when a user adds money to her transport card is the act of recharging (CargaTTP). In both cases, validation and recharging observations, the feature of interest is the transport card.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.47 Time dependencies in CRS definitions](#)



## 4.15 Combining spatial RDF data for integrated querying in a triplestore

**Contributed by:** Matthew Perry (Oracle)

### ? Full use case description (click to expand):

RDF datasets with spatial components are becoming more common on the Web. Many applications could benefit from the ability to combine, analyze and query this data with an RDF triplestore (or across triplestores with a federated query). Challenges arise, however, when trying to integrate such datasets.

- Inconsistent, incomplete, and non-standard metadata descriptions of spatial data
- Different representations of geometry data across different datasets
- Different spatial reference systems used across different datasets
- Different scales used across different datasets

For example, [Ordnance Survey](#) linked data uses British National Grid SRS and represents geometries with an Ordnance Survey-developed ontology, and [LinkedGeoData](#) uses WGS84 longitude latitude and represents geometries as GeoSPARQL WKT literals.

Best practices in the following areas would help make integration more straightforward.

- Agreed-upon vocabulary for metadata about spatial datasets
- Agreed-upon URIs for common spatial reference systems
- Recommendations for a default, canonical spatial reference system for spatial data published on the Web
- Recommendations for serializing geometries as RDF

Consistent metadata descriptions about spatial datasets will take out a lot of guess work when combining datasets, and standard URIs for spatial reference systems will be an important part of this metadata description. A recommended canonical SRS would make combining datasets more efficient by eliminating the coordinate transformation step, which would make federated GeoSPARQL queries more feasible.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.8 Default CRS](#), [5.12 Encoding for vector geometry](#), [5.40 Spatial metadata](#)

## 4.16 Dutch Base Registry

**Contributed by:** Linda van den Brink

### ? Full use case description (click to expand):

Dutch government data is for a large part open data. However, at the moment the data is difficult to find, and it cannot be easily linked to other data. It is not helping that most registrations are making use of their heavy backoffice standards for opening their data. These problems are counteracted by copying data of others, involving heavy expenses for collecting, converting and synchronizing the data, or by building expensive national provisions. The result is an abundance of copies and much doubt regarding the authenticity of the information.

A better solution would be to make the authentic data permanently available as linked data, so that everyone can use it and the datasets can be interlinked, resulting in more coherence and improved traceability and no more need for copying and synchronization.

There are now 13 Dutch 'base registries' containing government reference data: e.g.

- BAG (buildings and addresses)

- NHR (businesses, organizations)
- BRP (citizens)
- BGT (large scale topography)
- BRT (small scale topography)
- BRK (cadastre)
- WOZ (building tax)

A lot of these have geospatial content or refer to geospatial resources in other base registries (such as addresses). At the moment these references are informal and often incorrect or outdated. This means there is a need to express geospatial content in linked data (i.e. a standard vocabulary) and to perform spatial queries over linked data.

Geometries, especially lines and polygons, may contain many coordinates. For example, a municipal boundary could easily contain more than 1500 coordinate pairs. Compared to non-geometric properties, this can result in large amounts of data to transfer and process. The coordinates can easily be 95% of all data of an object when using polygons. The question rises whether there is a need for performance optimization and/or compression techniques for large amounts of coordinates. If so, there could also be a need to standardize such a technique, similar to the PNG format for encoding images.

Coordinate reference systems (CRS) are to geo-information what character encodings are to text. If you don't know which CRS is used, you can't use the coordinates. Different CRSs exist for a reason: localized CRSs provide more precise coordinates for a certain part of the globe. It is not possible for a global CRS to be as precise, for example because the continental plates move a few centimeters every year. For large scale data and applications this continental drift could be very relevant over time. Take for example the boundary of cadastral parcels. If this drift is not taken into account, there could be issues if parcel boundaries that were established e.g. 10 years ago are overlaid over recently acquired aerial imagery with high accuracy (e.g. 10 cm). There could be visual differences, while the actual situation did not change.

While the possibility to use different CRSs hinders interoperability (datasets using different CRSs cannot be easily combined, a complex transformation is necessary), on the other hand this option is perhaps needed for use cases where a high precision of coordinates is important. The BGT (large scale topography) is an example of such a use case.

A prototype application, based on linked data, where BGT, BAG, NHR and WOZ data is combined, is [here](#). BAG linked data is [here](#) ("Begrippen": the vocabulary; "BAG Data": the data)

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.3 Compressible](#), [5.19 Linkability](#)

## 4.17 Publishing Cultural heritage data

**Contributed by:** Lars G. Svensson

**? Full use case description (click to expand):**

Cultural Heritage Data such as library authority files are increasingly being published as Linked (Open) Data. Those datasets contain, among other entities, descriptions of spatio-temporal events such as *World War I*, the *Birth of Albert Einstein* (date and place) or *Martin Luther's pinning of his 95 theses*. The problem is that most of the spatio-temporal information is inexact. This inexactness ranges from time spans (*second quarter of the 9th century*, e.g. approx. 825-850, which also could be 823 or 852) to geographic entities such as *Renaissance Italy* (what did Italy look like at that time, and to what extent is the Italian renaissance as a time period different from the English?).

When cultural heritage institutions put their data on the Web, the staff members mapping their data to Web standards often do not have expertise in temporal or geospatial data formats. Formats such as

WKT are standardized but it is difficult to know if the mapping from the local data source is done correctly. A validator for commonly used formats such as WKT or GeoJSON would prove helpful.

Challenges include:

1. There is no framework available to describe fuzzy temporal information. ISO 8601 and the related XSD datatypes assume a precision that cannot be supplied by the data publishers.
2. In addition to that, OWL reasoning over cultural heritage datasets is severely hampered since OWL only accepts the datatypes `xsd:dateTime` and `xsd:dateTimeStamp` as temporal datatypes (cf. [Time Instants](#)).
3. Little or no possibilities to perform validation of temporal and/or spatial data formats (e.g. WKT).

Note: This use case has similarities to [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.4 Coverage temporal extent](#), [5.7 Date, time and duration](#), [5.15 Georeferenced spatial data](#), [5.28 Nominal temporal references](#), [5.29 Non-geographic reference system](#), [5.53 Update datatypes in OWL Time](#), [5.39 Space-time multi-scale](#), [5.49 Temporal vagueness](#), [5.43 Spatial vagueness](#), [5.55 Validation](#)

## 4.18 Dissemination of 3D geological data

**Contributed by:** Rachel Heaven

**? Full use case description (click to expand):**

The British Geological Survey (BGS), like all Geological Survey Organizations (GSOs), has as one of its principal roles to be the primary provider of geoscience information within its national territory. Increasingly the information provided is digital and dissemination is over the internet, and there is a trend towards making more information freely available. For BGS's 2D information this requirement has been met by the provision of [various OGC Web map services](#).

However, geological units and structures are three-dimensional bodies and their traditional depiction on two-dimensional geological maps leads to a loss of information and the requirement of quite a high level of geological understanding on the part of the user to interpret them. The geoscience data user community includes scientific users, but also includes many other stakeholders such as exploration companies, civil engineers, local authority planners, as well as the general public. It is therefore the aim of many Geological Surveys, including BGS, to move towards the provision of geological information as spatial 3D datasets on the Web that are accessible and useable by non-experts.

We have implemented a few ways of disseminating the 3D data on the Web (<http://www.bgs.ac.uk/services/3Dgeology/lithoframeSamples.html>, <http://www.bgs.ac.uk/services/3Dgeology/virtualBoreholeViewer.html>, <http://earthserver.bgs.ac.uk/>, investigation of augmented reality smartphone application) but the remaining issues are

- user should not need any special software for discovery and assessment purposes
- user can subset the data by x,y,z limits
- user can overlay their own area/volume of interest on the model (e.g. for tunneling projects)
- user may have to overlay their own (or an open data) DTM if the model is built using a non-open data DTM
- to represent complex geology (folding, faulting, intrusions etc) and varying data resolution the delivery method must be able to handle triangulated irregular networks (TINs) and heterogeneous resolution voxel grids.

If there existed a best practice or standard way of publishing this sort of data then it would encourage development of applications on the Web to handle them. The datasets that BGS is generating are:

- 3D spatial models of rock type and/or lithology currently represented as stacks of 2d subsurface grids representing the interfaces between different layers.
- 3D voxel datasets of x,y,z grid cell and geological, physical or chemical property (+ associated uncertainty).
- 3D LIDAR point cloud data e.g. of retreating cliff faces, underground mine structures.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.10 Discoverability](#), [5.15 Georeferenced spatial data](#), [5.23 Model reuse](#), [5.26 Multiple types of coverage](#), [5.31 Observed property in coverage](#), [5.33 Quality metadata](#), [5.34 Reference data chunks](#), [5.46 Support for 3D](#), [5.54 Use in computational models](#), [5.40 Spatial metadata](#), [5.50 Support for tiling](#), [5.3 Compressible](#), [5.19 Linkability](#)

#### 4.19 Publication of raw subsurface monitoring data

**Contributed by:** Rachel Heaven

##### ? Full use case description (click to expand):

The UK Government is funding a new Energy Security and Innovation Observing System for the Subsurface (ESIOS). ESIOS will consist of a group of science research facilities where new subsurface activities such as fracking (hydraulic fracturing) for shale gas can be tested and monitored under controlled conditions. This research will address many of the environmental issues that need to be answered for the development of the UK's home-grown, secure energy solutions. This includes carbon capture and storage, geothermal energy, nuclear waste disposal, underground coal gasification and underground gas storage.

Data will be collected from monitoring boreholes and from surface, airborne and satellite sensors. The raw scientific data will be published freely online, possibly in real-time or near real-time, to encourage transparency and public confidence in the industry, and to provide underpinning science for regulation.

The scientific data will consist of:

- location data for the sensor facilities;
- time series data consisting of x,y,z (depth), time, and an observed property (e.g. porosity, resistivity, density, methane content, fracture orientation, groundwater flow direction, seismic ground motion).

We want to be able to publish this raw data on the Web in such a way that it can be easily consumed by third party Web applications for visualization, spatio-temporal filtering, statistical analysis and alerts.

(Another similar use case is for publication of geomagnetic monitoring data, for which the primary outputs are time series tables or graphs, and the location data for the observation station is relatively simple and low accuracy)

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.16 Humans as sensors](#), [5.6 CRS definition](#), [5.11 Dynamic sensor data](#), [5.15 Georeferenced spatial data](#), [5.19 Linkability](#), [5.22 4D model of space-time](#), [5.23 Model reuse](#), [5.27 Nominal observations](#), [5.30 Observation aggregations](#), [5.35 Reference external vocabularies](#), [5.36 Sampling topology](#), [5.38 Sensing procedure](#), [5.37 Sensor metadata](#), [5.40 Spatial metadata](#),

[5.43 Spatial vagueness](#), [5.46 Support for 3D](#), [5.39 Space-time multi-scale](#), [5.51 Time series](#), [5.52 Uncertainty in observations](#), [5.57 Virtual observations](#)

## 4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches

**Contributed by:** Rachel Heaven

### ? Full use case description (click to expand):

BGS has valuable geoscience data in text form dating back 180 years, which is gradually being scanned and OCR to make it more accessible and searchable. Much of the information in the reports concerns observations or interpretations made at a location or about a named geological feature. We would like the relevant information in those documents to be retrievable from a Web search using coordinate limit criteria or using a place name criteria, so this use case requires a place name ontology (or federated ontologies).

For BGS's purposes the ontology should contain historical place names, named subsurface geological features (e.g. Widmerpool Gulf), palaeogeographic place names, and named submarine features (e.g. [GEBCO undersea feature names](#)).

To extend the capabilities into the vertical dimension then the ontology should also contain the names of qualitative earth realms (vertical divisions within the atmosphere, ocean and solid earth – such as in [the SWEET ontology](#)).

To extend the capabilities into the time dimension then the ontology should also contain the names of historical and geological time periods (e.g. [https://www.seegrid.csiro.au/wiki/CGIModel/GeologicTime#GeologicTime\\_XML](https://www.seegrid.csiro.au/wiki/CGIModel/GeologicTime#GeologicTime_XML), used in a recent example of geological age name parser at <http://www.agenames.org>)

With a resource like this, all text resources could be parsed to locate them in time and space.

Each named feature should have a spatial attribute, either as topological relations to other named features, or as spatial-temporal extent appropriate for various scales and with appropriate uncertainties (i.e. fuzzy definitions of geometry and time periods). Versioning will be important e.g. for administrative boundaries that change frequently.

(NB [Geonames](#) goes much of the way to meeting this requirement)

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.9 Different time models](#), [5.7 Date, time and duration](#), [5.46 Support for 3D](#), [5.28 Nominal temporal references](#), [5.48 Temporal reference system](#), [5.49 Temporal vagueness](#), [5.43 Spatial vagueness](#), [5.41 Spatial relationships](#), [5.56 Valid time](#), [5.40 Spatial metadata](#)

## 4.21 Driving to work in the snow

**Contributed by:** Cory Henson (Bosch RTC)

### ? Full use case description (click to expand):

Consider the following wildly-fictional scenario: Sue is driving to work in the snow on a cold Pittsburgh morning. On entry, her car recognizes the cold weather and automatically heats the interior to Sue's preferred temperature and turns on the defrost to clear the frost from the wind-shield. On the way, the snow causes significant traffic delay on her route forcing her to re-schedule an early morning meeting. In response, the car suggests she stop at her favorite nearby coffee shop for a Flat White until the roads are clear.

The scenario above requires access to multiple types of observation, spatial, and temporal data which may be local to the car or available on the Web.

The different types of observation data may include:

- sensed environmental observations (e.g., freezing-temperature, frost-on-windshield);
- user behavior observations (e.g., Sue sets the thermostat to 75F, Sue stops at the coffee shop, Sue updates her calendar);
- observations on the Web (e.g., current weather conditions, traffic conditions).

The different types of spatial data may include:

- points-of-interest (e.g., home, work, coffee shop);
- spatial relations (e.g., coffee shop is NEAR Sue's current location).

The different types of temporal data may include:

- calendar events (e.g., Sue's meeting on calendar);
- schedules (e.g., schedule when the coffee-shop is open);
- temporal relations (e.g., the predicted time to reach work occurs after the beginning of a scheduled meeting).

In addition, the car uses light-weight communication protocols, such as CoAP and/or MQTT, to exchange data (i.e., observations, spatial, and temporal data) between networked components.

This scenario may face the following general challenges for representing, managing, and querying observation data:

- discover observation streams within a given spatial context (and matching some criteria);
- query for observations within a given spatiotemporal context (and matching some criteria);
- publish and subscribe to an observation stream within a given spatial context (and matching some criteria);
- filter an observation stream based on given set of criteria (e.g., quality, feature-of-interest, observed-property);
- represent data with a small syntactic footprint (for exchange between resource-constrained devices);
- integration of SSN with existing, well-known, IoT protocols (CoAP, MQTT, etc.);
- represent "high-level" qualitative observations (such as observations of user behavior, e.g., Sue stopped at the coffee shop).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.16 Humans as sensors](#), [5.7 Date, time and duration](#), [5.10 Discoverability](#), [5.11 Dynamic sensor data](#), [5.15 Georeferenced spatial data](#), [5.18 Lightweight API](#), [5.24 Moving features](#), [5.27 Nominal observations](#), [5.28 Nominal temporal references](#), [5.37 Sensor metadata](#), [5.20 Machine to machine](#), [5.3 Compressible](#)

## 4.22 Intelligent Transportation System

**Contributed by:** Antoine Zimmermann

**? Full use case description (click to expand):**

Intelligent Transportation Systems (ITS) can be defined as the application of advanced information and communications technology to surface transportation in order to achieve enhanced safety and mobility while reducing the environmental impact of transportation. The addition of wireless communications offers a powerful and transformative opportunity to establish transportation



connectivity that further enables cooperative systems and dynamic data exchange using a broad range of advanced systems and technologies. - See more at: [http://www.its.dot.gov/standards\\_strategic\\_plan](http://www.its.dot.gov/standards_strategic_plan)

ITS do so by exploiting information from various origins, especially the Web. Several Web sites, Web services, datasets related to public transport services, traffic, road network structure, localized incidents, and so on have to be exploited in order to convey users with the best decisions, in real time. All of these information sources rely considerably on spatio-temporal data. The challenge is to model dependency in both space and time seamlessly and simultaneously so that the accuracy of analysis can be improved (for instance for regulation of multimodal transportation network) or the processing of aggregated information can be simplified (for instance for multimodal traveler information system). For such systems to work well in a pervasive way, it is also important that the system can easily discover relevant datasets/services based on the user current location.

Consequently, such systems would greatly benefit from standardized formats, standard practices for publishing or making spatial data available online. A standard way of indicating time and time frames would also help correlate several temporally situated entities such as bus schedules, real time bus position, and traffic light durations.

A specific type of ITS is Advanced Traveler Information System (ATIS) that computes an accurate travel duration. To do so, ATIS should be able to combine data following different spatio-temporal scales. These data could be related to the current or anticipated network traffic (the congestion level), the network topology (number of lines on the routes, presence of traffic lights, etc.), the presence of expected events (which can be static, like work on the road, or dynamic like demonstrations), the weather state (the presence of rain or mist on a part of the itinerary), and so on.

ITSs can also be dedicated to the management of parking spots. For a driver, the choice of its parking spot is a multi-criteria decision process that takes into account static data given by the description of the infrastructure (whether the spot is private or public, reserved for disabled, etc.), dynamic data given by sensors (like traffic) and personal data (destination, budget).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.21 Mobile sensors](#), [5.24 Moving features](#), [5.27 Nominal observations](#), [5.39 Space-time multi-scale](#), [5.48 Temporal reference system](#)

#### 4.23 Optimizing energy consumption, production, sales and purchases in Smart Grids

**Contributed by:** Antoine Zimmermann

##### ? Full use case description (click to expand):

In smart grids, energy management has to take into account the fact that any energy consumer may also be a producer (they are thus "prosumers"). This leads to new load balancing problems as well as new forms of economic exchanges regarding selling and buying energy. In order to take an informed decision on what amount of energy to buy from whom at what cost, or to sell, decision algorithms must use information that can be local (their own consumption and production), global (statistical data on seasonal household energy consumption), and possibly external to grid (meteorological data). In such context, reliance on Web data from several sources adds real value to the decision process.

The kind of data that has to be considered are, for the most part, highly fluctuating: weather for assessing heating needs, stock exchange for pricing appropriately, current and future offer and demand, etc.



There is a need for a temporal model that covers historical data for statistical analysis, short term timestamped sensed data, and data about future predictions.

The need for spatio-temporal information is even increased if the smart grid is including electric vehicles that can serve as energy producers when they are not consuming electricity for recharging.

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related requirements:** [5.19 Linkability](#)

#### 4.24 Linked Data for tax assessment

**Contributed by:** Luigi Selmi (via [public-sdw-comments](#))

##### ? Full use case description (click to expand):

Tax assessments are based on the comparison of what is due by a citizen in a year for her ownership of real estates in the area administered by a municipality and what has been paid. The tax amount is regulated by laws and based on many criteria like the size of the real estate, the area in which it is located, its type: house, office, farm, factory and others. Taxpayers can save money from the original due depending on the usage of the estate. A family that owns the house in which they live can save the entire amount. Many other regulations lighten in different ways the burden of the tax for other categories of taxpayers. Furthermore the situation about a taxpayer changes over the years in relation to her properties share and family status. Due to the many different situations met, an employee in charge of performing tax assessments on behalf of a municipality must collect many information before being able to assert with a good degree of confidence that a difference between the original amount and what has been paid is not justified and an advice has to be sent to the taxpayer starting a long and expensive process to recover the difference. Currently each single assessment requires the employee to collect information from different public administrations Web sites, archives, registries, documents. Data scattered in so many silos and formats dramatically reduce employees productivity and assessment effectiveness at the point that it is not always clear whether the money recovered is worth the cost of the assessment. A Linked Data approach for sharing spatial and temporal data would certainly increase the productivity of the assessor.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.10 Discoverability](#), [5.19 Linkability](#)

#### 4.25 Images, e.g. a time series of a water course

**Contributed by:** Kerry Taylor (on behalf of Jamie Baker, Australian Commonwealth Department of Communications)

##### ? Full use case description (click to expand):

This use case is provided to extend three primary use cases already before the Working Group:

- [4.7 Publishing geographical data](#)
- [4.8 Consuming geographical data in a Web application](#)
- [4.10 Publishing geospatial reference data](#)

More broadly the Commonwealth of Australia has developed a National Map Web platform which is currently making available authoritative national datasets. There is a need to recognize that data can an image (e.g. an image-based tile set for example) and therefore in itself also create a time series-based resource (for example, the change in a water course over time which can be visualized as time series layers). Indeed both the ISO and OGC have recognized this in their standards. It is the Australian Government Department of Communication's view, as the lead agency stewarding spatial

data policy, that image-based resources should also be included in the consideration of this working group as it relates to geographic and spatial features geometries. Wheresoever possible the Commonwealth view is to maintain the highest applicability of a standard or best practice guide and not limit conformance options for data holdings (especially of public origin). This also applies to cadastral and other data at the state/territory level which could show the change in land parcel, development or other property and built environment features over time. In terms of our future cities, sensors and other data sources may also need linkage to image-based resources for citizen use. As such:

1. image-based data needs to be recognized and discoverable as it too tells a story;
2. time series data can be visual and therefore visualized as a data layer or an image in time or both;
3. linking of data whether image-based or not should be discoverable and mix or match just as easily as data of a similar or dissimilar type;
4. metadata conformance is crucial to support discoverability, interoperability and provide information on the nature of a data resource (e.g. Time series image or time series data).
5. a challenge exists in maximizing the interoperability of geographic and/or spatial data both above and below ground level. (An example may be subsurface water imaging data combined with sensor data).

This additional information supports a broader need for SSN, Coverage and Time considerations for the above three current use cases.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.10 Discoverability](#), [5.15 Georeferenced spatial data](#), [5.26 Multiple types of coverage](#), [5.34 Reference data chunks](#), [5.39 Space-time multi-scale](#), [5.51 Time series](#), [5.46 Support for 3D](#), [5.37 Sensor metadata](#), [5.6 CRS definition](#), [5.35 Reference external vocabularies](#), [5.19 Linkability](#), [5.32 Provenance](#), [5.36 Sampling topology](#), [5.7 Date, time and duration](#), [5.33 Quality metadata](#), [5.56 Valid time](#)

#### 4.26 Droughts in geological complex environments where groundwater is important

**Contributed by:** Chris Little (on behalf of Andrew G Hughes, British Geological Survey)

##### ? Full use case description (click to expand):

During the period 2011-2013, the UK faced an extraordinary drought only to be saved from a likely major crisis by very high spring and summer rainfall (CEH, 2012). Government asked questions such as “how much water is left in the tank?”, at which point managers and regulators realized that they didn’t know the answer with any certainty (EA, 2012). A National Drought Group of leading stakeholders was formed and led by the Chief Executive of the Environment Agency (EA) reporting to the Secretary of State for Environment, Food and Rural Affairs. The whole of the UK was affected, but the south-east most severely because of the lower rainfall and high water demand.

When will London run out of water? How socially accepted is drought and associated water restrictions to the general public? When will water company groundwater sources begin to switch off? Questions like these are critical for drought management, but often the science is not sufficiently advanced to address them, and where it is, not fully integrated to satisfactorily answer them.

The River Thames Basin is home to 13 Million people, considerable industry and valuable aquatic ecosystems all of which require the effective and sustainable management of the water environment in the basin to thrive. A thorough understanding of the hydrology of the basin is vital to underpin this management to ensure the best use of resources. This is particularly important given the twin pressures of increasing water consumption and climate change. The groundwater system of the Thames consists of around 12 aquifers most of which are not hydraulically connected, except via the

River Thames and its tributaries. These aquifers are locally very important for water supply and their provision of base flow sustains the ecology of the river system in dry summer periods and droughts.

To properly simulate the system then a good geological understanding has to be translated into a hydrogeological knowledge and then simulated. This is not a straight-forward task. The important elements of the system include: 3D geological understanding encapsulated into a geological model, dynamic model of the surface processes, groundwater and river systems along with a model of water supply (e.g. IRAS; Mastrosov et al., 2010). This will ultimately involve:

- Static data (updated as understanding increases);
- Data passed in a one way chain between models runs separately;
- Dynamically linked models to investigate feedbacks between different parts of the system.

The aim would be to develop an integrated system that can address the question “When will water company groundwater sources begin to switch off” or similar.

## References

- CEH. (2012). Retrieved 9 Dec 2015, from [From drought to flood: dramatic turnaround in UK water resources](#)
- EA (2012). Review of the 2010-2012 drought and prospects for water resources in 2013. Environment Agency Report.
- Matrosov E.S., Harou, J.J. and Loucks D.P., 2010. A computationally efficient open-source water resource system simulator - Application to London and the Thames Basin. Environmental Modeling & Software. Volume 26, Issue 12, December 2011, Pages 1599–1610.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.19 Linkability](#), [5.32 Provenance](#), [5.33 Quality metadata](#), [5.27 Nominal observations](#), [5.46 Support for 3D](#), [5.57 Virtual observations](#), [5.11 Dynamic sensor data](#), [5.54 Use in computational models](#), [5.51 Time series](#)

## 4.27 Soil data applications

**Contributed by:** Simon Cox (on behalf of Peter Wilson, Bruce Simons @ CSIRO)

### ? Full use case description (click to expand):

1. Farmer wants to know soil types and properties in an adjacent property which he is considering purchasing. ... the global soil community wants to be able to publish standardized soil site description/classification and soil analysis data which can then be collated (such as in the Harmonized World Soil Database) and used for modeling the spatial distribution of soil properties (digital soil mapping) (for example as part of the ISRIC 1kmSoilGrids or IUSS GlobalSoilMap projects). This would utilize our efforts with the SoilML (ISO28258) meta-model and the fully attributed implementable extension (currently ANZSoilML) to deliver the standardized soil data services.
2. Compare similar soils and productivity responses to addition ... the IUSS GlobalSoilMap project wants to deliver high resolution estimates of functional soil properties as 90m raster data across the globe, constructed through bottom up, country delivery of standardized data. Our TERN Soil and Landscape Grid of Australia project has now developed the required data for Australia and we need to progress delivery as standardized services (using WMS, WFS and WCS) as a demonstrator of how to do this globally. The GlobalSoilMap project (maybe through the UN Global Soil Partnership, or via ISRIC the World Soil Data Centre) could develop a portal to allow discovery/display of all the contributing country data and to facilitate user (e.g. global climate or food security modelers) connection to the services from multiple sources.

3. Agronomist needs observations of soil properties such as pH, Nitrogen, and soil-type and derive new properties using locally defined pedo-transfer functions ... Access observed and interpreted soil properties (by soil type and/or by spatial distribution) in a standard format that allows using algorithms/pedotransfer functions to use these properties to calculate additional interpreted soil properties.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.10 Discoverability](#), [5.23 Model reuse](#), [5.26 Multiple types of coverage](#), [5.30 Observation aggregations](#), [5.31 Observed property in coverage](#), [5.33 Quality metadata](#), [5.35 Reference external vocabularies](#), [5.39 Space-time multi-scale](#), [5.57 Virtual observations](#), [5.38 Sensing procedure](#), [5.16 Humans as sensors](#)

## 4.28 Bushfire response coordination centre

**Contributed by:** Simon Cox (on behalf of Paul Box, Simon Cox and Ryan Fraser @ CSIRO)

### ? Full use case description (click to expand):

This user story covers a number of interrelated use cases:

- UC1 Search for place name
- UC2 Disambiguate place, based on type and provide
- UC3 Find the location of a place (centroid, representative location)– where is it?
- UC4 Find information about a place
- UC5 Find and use a suitable alternative geometric representations of a place – its shape, for a given scale and application
- UC6 Discover, access and use measurements relating to a place
- UC7.1 Discover related places - find synonyms
- UC7.2 Discover related places - topologically related places such as neighboring, or containing regions

### User story

A bush-fire is underway in Blue Mountains NSW. An incident management team has been established, with an incident controller from NSW Fire and Rescue in charge, based in a coordination centre. She notices that the Twitter analytics feed has flagged a Tweet that mentions a fire in 'Springwood'. The location of the Tweet shows on the incident map as being in the 'Blue Mountains' – (this location is a geocode of place name used in the Tweeter's Profile). The watch officer enters 'Springwood' [uc1] in a multi-gazetteer index and gets 41 hits for Springwood in Australia. She zooms to the Blue Mountains and sees there are 13 places called Springwood. These results include a school, a number of rural properties, and an official suburb named Springwood [uc2] near which most of the other places are located. The officer selects the suburb name in the National Gazetteer, and the map zooms to the bounding box for the selected place [uc3]. The controller wants to find more detailed information about the place and clicks on its identifier (a URI for the place). The user is provided with information about:

- Identity of the place and data sets/collection to which it belongs [uc4]– e.g. the national gazetteer.
- Available geometries for the place [uc5]- e.g. a feature-service for a gazetteer data source, hosted by Geosciences Australia.
- Existence of links to additional sources of information about the place [uc6] – e.g. links to data about the suburb provided by the local government.
- Synonyms for the place – alternative identifiers used for the place [uc7.1] – e.g. identifiers used for suburbs in the Australian Statistical Geographic Standard (ASGS).
- Related places [uc7.2] – e.g. a local government area that contains the suburb, or hydrological reporting units cover the suburb.
- information about related places e.g.:

- population profiles for the place codes using ASGS codes [uc7.1&6]
- nearby sensors providing hydro/met data that are in the containing hydrological reporting unit [7.2 & 6]

An official report of the fire location now also appears on the screen, from a fire service operator on the ground. With the fire location now confirmed, the priority is to identify the areas at risk, and assess community impact, locate possible evacuation centers and set in place evacuation plans.

A fire spread model has been run using reported fire location. An impact analysis and evacuation plan is being developed. The predicted fire spread polygon is shown on the incident map. To assess affected population the analyst needs to find population geography and data. [uc5]

The Springwood place landing page has a link to get information about the data source [uc4]. This is the national gazetteer, which provides bounding box geometries only. This is not accurate enough for her analysis. Returning to the graph of linked resources, the analysis determines that the gazetteer entry has a synonym in the ASGS [uc7.1]. She clicks on this and sees it's the census geography for the same place (Springwood suburb) [uc4]. She finds the data source link which she discovers uses polygonal geometry and adds this as a service to her map.

She is now able to see that Springwood suburb will be within the predicted fire polygon and will need to be evacuated. The analyst clicks linked information resources for Springwood (the synonym in the ASGS dataset) and a graph of related resources is displayed [uc 7.1, 7.2, 6]. This shows a link to a suburb profile based on the most recent national population and census data is available the analyst clicks on this and a query for predefined measurements (plus the place identifier) is passed to a census data cube service. The analyst visualizes and saves the result set locally.

The analyst also notices that information resources are connected to Springwood (in the gazetteer dataset) [uc 7.1, 7.2, 6]. These include a link to a containing local government area (LGA) and links to LGA emergency management resources on the Website. This lists contact information and evacuation centers information which is used to inform development of the evacuation plan.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.19 Linkability](#), [5.37 Sensor metadata](#), [5.43 Spatial vagueness](#)

## 4.29 Observations on geological samples

**Contributed by:** Simon Cox

**? Full use case description (click to expand):**

Geological samples are retrieved from the field and then processed in the laboratory to determine various properties, including chemistry, mineralogy, age, and petrophysical properties like density, porosity, permeability.

Samples obtained as part of economic activities, such as mineral exploration, are usually processed in commercial assay and chemistry labs. For QA/QC purposes, each batch of samples will have a number of control samples inserted, for which the concentration of particular chemical species are already known. For confidentiality reasons the location information associated with each sample is not provided to the lab, but must be re-attached during the interpretation phase. During processing, many derived samples will be generated by various physical and chemical procedures. In some cases the derived samples are strict sub-samples, whose intensive properties are intended to be the same as the parent. In other cases, the split is 'biased', with the derived sample intended to select a specific sub-sample, defined by a particular particle size, density, magnetic properties, etc. The link from the derived sample to the parent sample must be preserved, and the link from the parent to the location from which it was obtained also. In some cases the location is associated with another



sampling artifact, such as a drill-hole or traverse or cruise, with the latter carrying the detailed location information.

In a research context some samples have a particularly high-value, having been obtained by an expensive process (involving drilling or ships or spacecraft) or from a location that is hard to visit (remote, offshore, in space). These samples are sometimes sub-divided and distributed to multiple research teams or labs for different specialized observations. Each lab will run its own LIMS system, which will usually assign a local identifier for the sample. When the results of these observations are reported, it is necessary that observations from different labs can be correlated with each other, so that the complete picture around each sample can be assembled.

These stories focus on sensing applications involving ex-situ sampling, where a location is visited and a specimen obtained using some sampling process, then transported to one or more laboratories where it is processed into one or more sub-samples and various observations made. Sample identity is usually key, and the relationships between samples, between samples and other artifacts of the sampling process, and also with other geographic features or locations. The sampling time and analysis and reporting time are all different.

Similar process apply to botanical sampling, and to environmental sampling (water, air, dust).

**Related deliverables:** [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.9 Different time models](#), [5.15 Georeferenced spatial data](#), [5.19 Linkability](#), [5.23 Model reuse](#), [5.32 Provenance](#), [5.35 Reference external vocabularies](#), [5.37 Sensor metadata](#), [5.13 Ex-situ sampling](#), [5.36 Sampling topology](#), [5.38 Sensing procedure](#), [5.16 Humans as sensors](#), [5.48 Temporal reference system](#)

## 4.30 Spatial sampling

**Contributed by:** Simon Cox

**? Full use case description (click to expand):**

Most observations are made on *samples* that are selected to be somehow representative of the feature of ultimate interest which is being characterized. The sample may be statistical, or related to accessibility, or may be of a proxy phenomenon which can be related to the property of ultimate interest.

In environmental observations, certain spatial distributions are commonly used across multiple disciplines, such as

- grids
- transects, cruises, flight-lines
- cross-sections
- physical sampling

These may be related to each other (samples taken along a drill-hole, stations on a traverse, shots along a seismic section) and the relationships provide a kind of 'topology' of sampling which assists access and processing. Sampling features may also be associated with other organizational structures, such as cruises, field trips, campaigns, projects, missions, orbits, deployments, platforms, which are used for discovery and for navigation within a datasets. Sampling strategies are often combined with an observational procedure or instrument to define a standard 'protocol' for observations. The protocol may be identified by name.

*Additional comment by Rachel Heaven:*

The locations within a spatial distribution may have a different degree of certainty with respect to each other than the positional certainty of the spatial distribution as a whole e.g. the ends of a

sampling traverse may be known in a national or global coordinate reference system to +/-5m accuracy; soil samples may be taken along the traverse at every 1m +/-0.01m interval, or species sightings that need to be re-visited may be described in real world terms such as "half way up the burn on the left hand bank". Similarly, measurements taken down a drill hole are known accurately in down hole depth with respect to the drilling datum, but less accurately in true vertical depth and with respect to a national vertical datum. If the hole deviates from vertical, the absolute location will be even more uncertain.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.6 CRS definition](#), [5.15 Georeferenced spatial data](#), [5.19 Linkability](#), [5.21 Mobile sensors](#), [5.23 Model reuse](#), [5.37 Sensor metadata](#), [5.39 Space-time multi-scale](#), [5.43 Spatial vagueness](#), [5.36 Sampling topology](#), [5.52 Uncertainty in observations](#), [5.40 Spatial metadata](#)

#### 4.31 Select hierarchical geographical regions for use in data analysis or visualisation

**Contributed by:** Bill Roberts (based on needs arising from Swirri's own work)

##### ? Full use case description (click to expand):

Statistical data is frequently referenced against geographical regions. A common requirement is to select a collection of non-overlapping regions with complete coverage of a 'parent' region (a 'Mutually Exclusive Collectively Exhaustive' - MECE - set). This might be used for data aggregation: given the population statistics for all council areas in Scotland, calculate the population of Scotland.

Or it might be for data visualization: retrieve data on average house prices for all parliamentary constituencies in the UK, then combine this with polygons of the constituency boundaries and use it to draw a choropleth map.

Although geographical data frequently includes 'contains' or 'within' relationships between larger and smaller areas, this is not sufficient for the above use cases. A larger area can be broken down into smaller areas in a variety of ways. Sometimes a combination of 'contains' relationships and knowledge of the 'type' of region can be enough: it may allow separating out a particular level in a geographical hierarchy. But that doesn't allow for cases where regions of a particular type don't have full coverage of the parent. An example is 'parishes' in England. The collection of all current parishes is non-overlapping, but it is not exhaustive. Some locations are not in any parish.

The variation of administrative, statistical and political geography over time is also an issue. A particular division of a region into sub-regions may be valid only for a specific period. For example, council boundaries change from time to time. At any given time, there is a MECE set of council boundaries in a region, but the boundaries change occasionally, so if analyzing data on English councils from 2014, a different set of areas is required than would be needed for say 2012. 2012 might involve areas A,B,C,D,E,F whereas 2014 requires A,B,C,D,G,H. A similar issue arises when dividing Europe up into countries for example.

These are common problems and there are probably many separate solutions already in active use. It would be useful however to have a standardized vocabulary to represent these kind of relationships, which would increase the interoperability of data analysis tools.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.42 Spatial operators](#), [5.41 Spatial relationships](#), [5.56 Valid time](#)



## 4.32 Satellite data processing

**Contributed by:** Kerry Taylor (informed by Matt Paget and Juan Guerschman, CSIRO)

### ? Full use case description (click to expand):

*Vegetation fractional cover* is a key metric for monitoring land management, both in pastoral and agricultural settings. The aim is to estimate the fractions of photosynthetic and non-photosynthetic vegetation and the remaining fraction of bare soil. Fractional cover is computed from both Landsat and MODIS satellite surface reflectance products but calibration and validation via ground-based observations is also needed.

The method for computing fractional cover was developed by comparing geo-aligned Landsat sensor and two MODIS-derived surface reflectance products. Scaling issues associated with different sensors and spatial resolutions were addressed, along with locally-measured effects of soil color and soil moisture.

Source Data:

1. Data from the TM (Landsat 5) and ETM+ (Landsat 7) sensors were downloaded from the United States Geological Survey (USGS) and corrected for atmosphere, topographic and bi-directional surface reflectance effects, standardized to a nadir view (i.e., satellite zenith angle is zero) and a solar zenith angle of 45°
2. The MODIS 16-Day Nadir BRDF-Adjusted Reflectance product (MCD43A4) provides surface reflectance data as if they were taken from the nadir view and uses the solar angle calculated at local solar noon (Schaaf et al., 2002). MCD43A4 data were obtained from the Land Processes Distributed Active Archive Center (LPDAAC). These data were resampled to a geographic coordinate system.
3. The MODIS surface reflectance products from the Terra satellite (MOD09) provide an estimate of the surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering and absorption (Vermote et al, 2011; Vermote & Kotchenova, 2008). The 8-day composite (MOD09A1) selects the best possible observation from an 8-day window, selected on the basis of high observation coverage, low viewing angle, cloud free and low aerosol loading, with a spatial resolution of 500 meters and seven spectral bands. The MOD09A1 surface reflectance data used here were subject to the same geometric processing as the MCD43A4 data.
4. Field measurements of fractional cover were obtained at 913 sites across Australia (ABARES, 2013). Exactly 78 of those sites were sampled repeatedly, resulting in a total of 127 observations (a measurement taken at a given site on a specific date). The field observations did not systematically record soil color or soil moisture. Recent observations include soil color (Munsell Soil Color Charts, 1994), yet soil moisture is only subjectively measured (i.e., wet or dry). To derive these two soil properties consistently and objectively for the entire series we used information derived from other sources.

As a result, a combined product is proposed which gives flexibility to use MODIS-derived estimates when large areas and high temporal repetition is desired, and Landsat-derived estimates when high spatial resolution is essential and/or when data prior to 2000 is needed. The algorithms needed for implementing a fractional cover product based on a blended Landsat-MODIS product are given [\[1\]](#).

Now, this fractional cover coverage product over Australia is computed monthly and [distributed by the NSW government](#), where it is used for their [Dustwatch program](#) amongst other things. The Dustwatch program publishes monthly (PDF) reports of wind-related erosion and groundcover change.

**[1]** [Guerschman et al. "Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and](#)

[MODIS data", Remote Sensing of Environment, to appear 2015.](#) (Note that this paper provides several references to similar approaches to using multispectral coverages to determine vegetation).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.6 CRS definition](#), [5.15 Georeferenced spatial data](#), [5.21 Mobile sensors](#), [5.26 Multiple types of coverage](#), [5.32 Provenance](#), [5.37 Sensor metadata](#), [5.16 Humans as sensors](#), [5.57 Virtual observations](#), [5.38 Sensing procedure](#), [5.17 Independence on reference systems](#), [5.40 Spatial metadata](#), [7.1 Definition of SSN profiles and checking compliance against the model](#)

### 4.33 Marine observations - eMII

**Contributed by:** Simon Cox (on behalf of [IMOS eMII](#))

? Full use case description (click to expand):

#### Water column observations

I am a modeler and I want to find, filter and extract water column data that has been collected by profiling instruments (e.g. CTD, XBT, ARGO Floats, CTD's mounted on animals) so that I can prime my models. I want to easily be able to discover these data either by nominating the collection device type, or by nominating data parameters of interest. Once I can see which datasets meet my broad discovery criteria, and whilst still in the Portal, I want to be able to filter out data that is not of interest (e.g. those data outside of my region of interest, those not covering my features of interest, data which is unqualified, data below a certain depth and data from institutions I don't trust). I want to extract these data in a harmonized format (i.e. receive a single file of aggregated data expressed using common data fields, or in multiple files where each file has a similar syntactic and semantic encoding). I don't want to have to spend time transforming datasets with differing formats, nor guess which data fields in datasets from different sources are semantically covering the same information. I need to know what each field in the downloaded data represents.

#### Reconfiguring data ingestion process

I am an eMII Project Officer. I spend my day pulling data from partner services and transforming it so that it can be published through the 1-2-3 Portal. Just when I think I have tweaked all of the systems I need to in order to successfully ingest and publish provider data, the provider changes his/her data format, schema syntax or semantics. I then have to re-write or re-configure systems to obtain any new data. This happens very frequently. Even data providers who supply me with data from the same instruments use different data encodings and formats so I have to create individualized database tables to manage their incoming data. I would like data providers to agree on common schema for expressing similar data types and in collaboration develop some 'governance' rules surrounding data publication to the Portal to reduce the time I spend on repetitive (and often unnecessary) tasks.

#### Observations along profiles

I am an eMII Project Officer and I manage multiple profile type datasets (e.g.: Argo profile, seals profile, XBT profile ...). I want to be able to assess the quality of the data provided by partners before inclusion into the IMOS database and making it available through the IMOS portal. I want to set up a system whereby every new profile will be compared with existing profiles available in the IMOS database. The incoming profile will have to conform to a standard format so that it is relatively easy to implement and develop different set of rules to enable comparison with existing profiles. The system will send me an alert if one or multiple profiles failed some tests. Then I will be able to follow up more quickly with the corresponding partners to check if an error occurs during the processing of

the data or if actually the data is correct. In the end, this process will enable an increase of the quality of the data provided to the end user.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.10 Discoverability](#), [5.21 Mobile sensors](#), [5.23 Model reuse](#), [5.30 Observation aggregations](#), [5.31 Observed property in coverage](#), [5.33 Quality metadata](#), [5.37 Sensor metadata](#), [5.46 Support for 3D](#), [5.36 Sampling topology](#)

#### 4.34 Marine observations - data providers

**Contributed by:** Simon Cox (on behalf of [IMOS eMII](#))

? Full use case description (click to expand):

##### **Atlas of Living Australia (ALA)**

The ALA would like to integrate their data into the AODN. The ALA serves a range of Web services including WMS and corresponding ISO19115/MCP metadata. The ALA's use case is unusual in that it has tens of thousands of WMS layers and metadata records. This data cannot be added to version 2 of the AODN portal because it is too large to be harvested into the menu tree. ALA will need to be integrated with the 123 portal. Dave Martin's team has created a proof of concept integration. There would be a single metadata record for ALA, which will allow it to be discovered in Step 1 of the portal. After proceeding to Step 2 the user would see something like [this mockup](#).

##### **Reef Life Survey (RLS)**

RLS have visited 2500+ coral and rocky reef sites and have conducted approx 6000 surveys across those sites. Each survey is conducted at a nominal 'site'.

###### *Methods 1 & 2*

During a survey observations are recorded of each of (1) vertebrate species abundance and biomass and (2) invertebrate species abundance.

###### *Method 3*

During a survey downward looking photographs are taken. The photographs are sequenced 1-20 for each site and are not geolocated. Subsequently, the photographs are scored at 5 points within each image. Each point is scored into one of 36 categories. Parameters are date, depth, resolution, major category, minor category, numerical value.

###### *User Story 1*

I'm interested in ecosystems, reef assemblages and inter-species interactions. Show me what vertebrate and invertebrate species were found at this site (and any corresponding location/depth/habitat information).

What I want to see:

- By site, all vertebrate species abundance and biomass values (for each survey?)
- By site, all invertebrate abundance values (for each survey?)

## User Story 2

I have a large-scale question to ask about a particular species, for example, I am interested in broad distribution patterns of lobsters, plankton dispersal and settling rates. Show me all of the sites that were surveyed and the presence/absence of a particular species at each of those sites.

What I want to see:

- Across all sites that have been surveyed by RLS, the latitude, longitude, depth and presence/absence of a particular vertebrate or invertebrate species (no time component)
- A time selector that shows changing or stable presence/absence values for repeat surveys at the same site

### Institute for Marine and Antarctic Studies (IMAS)

The IMAS use case includes a number of data collections. The main requirements is a mechanism to easily install the necessary applications. Ideally the AODN will host the applications (GeoNetwork, Geoserver etc) in the NeCTAR Cloud. This cloud based infrastructure will be managed by the AODN, but IMAS will have administrator accounts on each application and will be responsible for data content.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.15 Georeferenced spatial data](#), [5.27 Nominal observations](#), [5.16 Humans as sensors](#), [5.46 Support for 3D](#), [5.52 Uncertainty in observations](#)

### 4.35 Marine observations - data consumers

**Contributed by:** Simon Cox (on behalf of [IMOS eMII](#))

? **Full use case description (click to expand):**

#### Download Temperature and Velocity Data from NSW Moorings

The user would like to download temperature and velocity data from NSW moorings without downloading large numbers of NetCDF files, and without needing many clicks.

(Based on feedback from Robin Robertson).

#### Download Glider Data

The user would like an easy way to download the calibrated glider data. The user does not want the data delivered manually via drop box, or to face the difficulty of downloading NetCDF files in the way they are currently provided.

(Based on feedback from Robin Robertson).

#### Download ANMN Timor South moorings data

The user would like to download the ANMN Timor South moorings data - without needing 160 clicks.

(Based on feedback from Rebecca Cowley).

**Download XBT data**

The user would like to download XBT data from the portal. Not just the metadata - but the actual data.

(Based on feedback from Rebecca Cowley).

**Download NRS data**

The user would like to be able to download NRS moorings data.

(Based on feedback by Peter Thompson)

**Understandable by scientists from other fields**

The user would like to be able to understand the portal even though she is from another field (e.g. genomics).

(Based on feedback from Levente Bodrossy).

**Filter moorings data by deployment and instrument type**

The user would like to be able to filter moorings data by deployment and instrument type.

(Based on feedback from Craig Steinberg)

**Download data as CSV**

The user would like to download sea surface temperature from the Bass Strait as a CSV file - not NetCDF

(Based on feedback from Andre Chiaradia)

**Argo float data in a bounding box in the Southern Ocean.**

The user would like to download argo data from a particular region in the Southern Ocean.

(Based on feedback from Esmee)

Also see [additional stories in un-edited emails](#).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.38 Sensing procedure](#)

## 4.36 Building information management and data sharing

**Contributed by:** Linda van den Brink (with thanks to Henk Schaap - Gobar)

**? Full use case description (click to expand):**

The Dutch organization [Rijkswaterstaat](#) is responsible for the practical execution of the public works and water management, including the construction and maintenance of waterways and roads, and

flood protection and prevention in the Netherlands. The following is a use case about sharing Rijkswaterstaat information to support building processes.

A part of a highway is in need of a new tarmac layer. This involves a lot of information sharing between the contractor (Rijkswaterstaat), the organization responsible for organizing the project and the organization responsible for carrying out the maintenance work. Building Information Management (BIM) is used for this and the data is published via a Webserver.

The information being exchanged is asset management information, i.e. the location of the roads, tunnels, bridges, aqueducts, railway lines, surrounding terrain and water bodies involved (geometric information). In addition, non-spatial information is important, like the layers of material the road consists of, when it was last maintained etc. The spatial information is not only shared as 2D geometries: (detailed) 3D information is also used in the building process. The area involved contains about 9.000 relevant physical objects (about 100 different object types) and about 2.000 spatial objects that relate to the traffic network, about which information is exchanged.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.46 Support for 3D](#), [5.41 Spatial relationships](#), [5.55 Validation](#)

#### 4.37 Landsat data services

**Contributed by:** Kerry Taylor (on behalf of Aaron Sedgmen of Geoscience Australia)

##### ? Full use case description (click to expand):

Geoscience Australia (GA) maintains an archive of Landsat 5, 7 and 8 satellite imagery over the Australian continent from 1986 to present. The Australian Reflectance Grid 25 (ARG25) dataset produced from the Landsat archive is a collection of approximately 184,000 individual Landsat 5 and Landsat 7 scenes processed to the ARG25 specification. The ARG25 data are available to the public as file downloads and OGC Web services. The ARG25 collection can be searched using an OGC Catalogue Service (CSW) containing ISO 19115 metadata for each scene. Note that the ARG25 data services are expected to be superseded by the Australian Geoscience Data Cube when it becomes publicly available.

The ARG25 data services were promoted for use at the 2013 and 2014 Australian GovHack events, a competition aimed at mainstream (i.e. largely non geoscience/geospatial specialist) Web and application developers, to mashup, reuse, and remix open government data. The predominant use case for the ARG25 data at GovHack has been to perform spatiotemporal searches on the catalogue for an area of interest, retrieve and subset/stitch scenes from the Web services, and provide an animated time sequence of the imagery to allow users to visualize changes in land cover over time, e.g. "show me how my town has grown over the last 10 years".

The relative complexity and richness of the OGC/ISO Web service APIs and data exchange formats presented a barrier to developers who were accustomed to lighter weight APIs and formats optimized for rapid integration and mashing up of data in mobile and browser based applications. As a result, uptake of the ARG25 data at GovHack was less than hoped for, and we see this as indicative of the mainstream Web app development community in the real world.

GA has had much success with OGC and ISO standards in the sharing of data with government, research and industry partners and clients who are power users of spatial data and savvy in the standards. The less traditional users who are not spatial data specialists are less likely to access GA's data delivered using OGC and ISO standards, and this is a section of the user community that can potentially apply the data to highly innovative and entrepreneurial uses.

GA does use proprietary and quasi-standard light weight data exchange formats (e.g. JSON) and Web APIs (e.g. ESRI RESTful API) for delivering some geospatial data, although, in accordance with



government policy, it is GA's preference to adopt standards when possible to maximize interoperability.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.10 Discoverability](#), [5.34 Reference data chunks](#), [5.51 Time series](#), [5.4 Coverage temporal extent](#), [5.50 Support for tiling](#)

#### 4.38 Metadata and search granularity

**Contributed by:** Kerry Taylor (on behalf of Aaron Sedgmen of Geoscience Australia)

##### ? Full use case description (click to expand):

Geoscience Australia (GA) collects and curates Australian national geoscience and geographic data sets, for use by government, industry and the community. Following is a typical requirement of a user of GA data - "A mineral exploration company is collating published geological data for their mining lease. They are particularly interested in sample locations where gold ppm values are above 0.1ppm, occurring in greenstone, located on or near a fault line, and aged around 2.6 Ga".

Users are able to perform structured searches for GA datasets at the collection level using catalogues of ISO 19115 metadata. The collection level metadata provides users with download links to the packaged datasets, and in limited instances, links for accessing the data via Web services and/or applications that provide visualization and GIS analysis capability.

To perform fine grained searches within the dataset, such as feature level searches (e.g. show me sample locations in the OZCHEM database with gold ppm > 0.1), users must download the packaged datasets to their local environment and use their own tools to search through the data. Web services providing data access, e.g. WFS, and Web applications with GIS capability served by GA, e.g. the Rock Properties Explorer app, can provide some capability for searching within data collections.

Usability of search and discovery systems would be enhanced by having standards that define the line between spatial data and metadata in the context of searches, and standard methodologies for searching across collection level and feature level data.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.9 Different time models](#), [5.13 Ex-situ sampling](#), [5.28 Nominal temporal references](#), [5.43 Spatial vagueness](#), [5.46 Support for 3D](#), [5.39 Space-time multi-scale](#), [5.48 Temporal reference system](#)

#### 4.39 Crowdsourced earthquake observation information

**Contributed by:** Kerry Taylor (on behalf of Aaron Sedgmen of Geoscience Australia)

##### ? Full use case description (click to expand):

Geoscience Australia (GA) collects earthquake observation information from the general public to improve its knowledge base of the impact of earthquakes across the Australian continent. An online HTML form is available on the GA Website for members of the public to report their experiences of earthquake events. Information collected includes the person's location (both geographic and within a building), time of event, perception of intensity and observed effects on the built environment. The information provided by the public is rated against the Modified Mercalli Intensity (MMI) Scale and is used to improve the accuracy of shake maps for earthquake prone regions of the country. This feeds



into GA's understanding of exposure of the Australian built environment to natural disaster, and is used for disaster mitigation purposes including the determination of minimum building codes for various regions of the country.

There is potential for GA to improve the efficiency by which it obtains earthquake observation information from the general public by leveraging popular social media services (such as Twitter), and adopting common standards and best practices for collecting crowdsourced information.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.27 Nominal observations](#), [5.25 Multilingual support](#), [5.16 Humans as sensors](#), [5.43 Spatial vagueness](#), [5.37 Sensor metadata](#), [5.15 Georeferenced spatial data](#), [5.30 Observation aggregations](#), [5.21 Mobile sensors](#), [5.46 Support for 3D](#), [5.11 Dynamic sensor data](#), [5.45 Streamable data](#), [5.40 Spatial metadata](#), [5.19 Linkability](#), [5.20 Machine to machine](#), [5.42 Spatial operators](#), [5.5 Crawability](#), [5.10 Discoverability](#), [5.12 Encoding for vector geometry](#)

#### 4.40 TCGA / microscopy imaging

**Contributed by:** Erich Bremer

**? Full use case description (click to expand):**

Studying the morphology of disease at the cellular and sub-cellular levels using high resolution tissue images is extremely important to help understand the nature of various cancers. [The Cancer Genome Atlas \(TCGA\)](#) contains over 32,000 de-identified whole-slide microscopy images (WSI) of over two dozen cancer types. These images can contain between 100K-1M nuclei each. Biomedical informatics researcher have developed (and continue to develop) software to automatically segment nuclei for study. The spatial features of each nucleus and groups of nuclei as it relates to other nuclei combined with other linked data such as other morphological features (crypts, ducts, etc.) and/or patient lab results are used in analyzing and categorizing tissues and patients into groups and in comparing such groupings to understand disease mechanisms in a particular cancer type as well as across cancer types.

Representing nuclear segmentations is often done with binary masks or through polygon representations (e.g., the use of Well Known Text (WKT) representations) and also by leveraging work from the Geospatial community. However, in the case of nuclear segmentations, coordinate systems are 2D & 3D Cartesian based. Although the majority of work in this area is 2D-based, a growing segment of microscopy is also 3D-based as the technology develops and become more sophisticated. As living tissue can change over time through growth, infection, cancer, damage, etc. (as well as its associated organism's various properties) it is important that spatial locations of features such as nuclear segmentation be also represented in a temporal aspect for proper comparisons.

Samples of TCGA WSI data can be viewed at: <http://cancer.digitalslidearchive.net>

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.4 Coverage temporal extent](#), [5.29 Non-geographic reference system](#), [5.34 Reference data chunks](#), [5.46 Support for 3D](#), [5.10 Discoverability](#), [5.40 Spatial metadata](#), [5.50 Support for tiling](#), [5.17 Independence on reference systems](#), [5.32 Provenance](#), [5.51 Time series](#), [5.42 Spatial operators](#)

#### 4.41 Crop yield estimation using multiple satellites

**Contributed by:** Kerry Taylor with Zheng-Shu Zhou, CSIRO

**? Full use case description (click to expand):**

Space-borne Synthetic Aperture Radar (SAR) observations are coming on line rapidly over the next few years. SAR signals are represented in 2D space but can be processed to point observations in 3D space, or to 3D triangulated surfaces using polarimetric and phase information. A compelling use case for SAR data is crop classification, i.e., identification of cereal crops (wheat, barley, rice etc.), through analysis of the radar backscatters and polarizations over the variable-height surface and volume of the crop.

Multispectral (e.g. Landsat or MODIS) satellite imagery has been widely used to estimate vegetation cover and growth rates. It seems that combining SAR derived crop-type observations with Landsat-derived vegetation estimations could be used to estimate crop yields throughout the growth cycle. In principle, access to such data is open, but in practice it is difficult to get hold of and difficult to process for anyone other than well-connected scientists in developed nations. How could this data be opened up to a bigger community to help solve this problem, perhaps aided by reference to local knowledge?

We might expect that this will be used to assist in logistics and market planning in developed countries, and for food security in developing and war-torn nations.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.4 Coverage temporal extent](#), [5.44 SSN-like representation](#), [5.46 Support for 3D](#), [5.57 Virtual observations](#), [5.37 Sensor metadata](#), [5.15 Georeferenced spatial data](#), [5.19 Linkability](#), [5.39 Space-time multi-scale](#), [5.52 Uncertainty in observations](#), [5.33 Quality metadata](#), [5.31 Observed property in coverage](#), [5.26 Multiple types of coverage](#), [5.34 Reference data chunks](#), [5.54 Use in computational models](#), [5.32 Provenance](#), [5.51 Time series](#), [5.6 CRS definition](#), [5.51 Time series](#), [5.42 Spatial operators](#)

#### 4.42 Geospatial extensions to domain-independent metadata schemas

**Contributed by:** Andrea Perego ([European Commission - Joint Research Centre](#))

**? Full use case description (click to expand):**

[DCAT-AP \(DCAT application profile for data portals in Europe\)](#) is a metadata interchange format, based on and compliant with the [W3C Data Catalog vocabulary \(DCAT\)](#), defining a minimum set of metadata elements to ensure cross-domain and cross-border interoperability across European data portals.

Work is under-way for a development of an extension aiming at defining a DCAT-AP compliant representation of geospatial metadata, referred to as [GeoDCAT-AP](#).

GeoDCAT-AP will build on preliminary work concerning the [alignment of INSPIRE metadata with DCAT-AP \(referred to as INSPIRE+DCAT-AP\)](#). The objective is to consolidate this work, and to extend it in order to cover the set of metadata elements corresponding to the union of the [INSPIRE metadata schema](#) and the core profile of ISO 19115.

Methodologically, GeoDCAT-AP will be designed based on the following requirements:

1. agree on mappings with metadata elements available in DCAT-AP (e.g., title, abstract, license, distributions)
2. define alignment with elements not available in DCAT-AP, by using, whenever possible, popular RDF vocabularies

The requirement for point (2) is to identify an RDF representation of elements missing in DCAT-AP that can be re-used, if relevant, in other domains. E.g., elements concerning data quality (e.g., conformity, topological consistency) and granularity (e.g., spatial / temporal resolution), two notions

not supported in DCAT-AP (nor in DCAT). This implies that, in many cases, there may be the need of relaxing the ISO 19115 specification whenever it may be a barrier to re-use.

Another key design principle is that GeoDCAT-AP is not meant to be a replacement of ISO 19115 / 19139, but an alternative representation, built on a set of harmonized mappings. For this reason, the principle is that any ISO 19115-conformant metadata record could be transformed into a GeoDCAT-AP one, but the opposite is not a requirement.

Issues to be addressed include the following:

- Recommendations on the use of HTTP URIs for geospatial data and metadata
- Recommendations on the representation of geometries, and for the specification of bounding boxes, centroids, etc.
- Recommendations on the representation of spatial coverage (not limited to geometries)
- Recommendations on the representation of temporal coverage
- Recommendations for the specification of spatial and temporal reference systems
- Recommendations for the specification of spatial and temporal resolution
- Recommendations on the representation of data quality and lineage (provenance)
- Support for multilingual metadata

Related use cases:

- [4.15 Combining spatial RDF data for integrated querying in a triplestore](#)
- [4.10 Publishing geospatial reference data](#)
- [4.43 Improving discovery of spatial data on the Web](#)

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.1 Bounding box and centroid](#), [5.5 Crawlability](#), [5.7 Date, time and duration](#), [5.9 Different time models](#), [5.10 Discoverability](#), [5.12 Encoding for vector geometry](#), [5.17 Independence on reference systems](#), [5.19 Linkability](#), [5.28 Nominal temporal references](#), [5.20 Machine to machine](#), [5.25 Multilingual support](#), [5.29 Non-geographic reference system](#), [5.32 Provenance](#), [5.33 Quality metadata](#), [5.40 Spatial metadata](#), [5.48 Temporal reference system](#), [5.43 Spatial vagueness](#), [5.49 Temporal vagueness](#), [5.51 Time series](#), [5.56 Valid time](#)

## 4.43 Improving discovery of spatial data on the Web

**Contributed by:** Andrea Perego ([European Commission - Joint Research Centre](#))

**? Full use case description (click to expand):**

*Use case based on [work presented at INSPIRE 2014](#) by Adam Iwaniak (Wroclaw University of Environmental and Life Sciences, Poland)*

Geoportals offer effective discovery functionalities for specialists. However, as in most data portals, using basic free text search is usually far from being a satisfactory experience. Actually, users (both non-experts and specialists), when searching for data, are frequently making use of popular search engines as a first step to get to the data they are looking for.

Improving free text search in (geo)data portals is unlikely to address this issue. Moreover, it would not help users who do not know in which portal(s) the relevant data are available. Users will keep on using in any case search engines for this purpose.

An option would be optimizing geoportals for Web discovery, by implementing consistently SEO (Search Engine Optimization) techniques. The advantages include (but are not limited to) the following:

- Increase visibility of spatial data - as well as of the geoportals giving access to them.

- Enabling queries not limited to free text fields (e.g., dataset title and description), but concerning also coordinates, spatial / temporal resolution, etc.
- Enabling users to find the data better satisfying their requirements (in terms of spatial and temporal coverage, granularity, formats, access and licensing conditions, etc.)

Best practices for publishing geospatial metadata on the Web should be recommended. This includes, for example, standard serialisations of geospatial metadata in formats and vocabularies used by search engines to index Web resources - e.g., RDFa, Microdata, Microformats

Related use cases:

- [4.8 Consuming geographical data in a Web application](#)
- [4.42 Geospatial extensions to domain-independent metadata schemas](#)

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.5 Crawlability](#), [5.10 Discoverability](#), [5.12 Encoding for vector geometry](#), [5.17 Independence on reference systems](#), [5.19 Linkability](#), [5.19 Linkability](#), [5.20 Machine to machine](#), [5.25 Multilingual support](#), [5.28 Nominal temporal references](#), [5.29 Non-geographic reference system](#), [5.32 Provenance](#), [5.33 Quality metadata](#), [5.40 Spatial metadata](#), [5.43 Spatial vagueness](#), [5.49 Temporal vagueness](#)

#### 4.44 INSPIRE compliance using Web standards

**Contributed by:** Erwin Folmer, [Dutch Cadastre](#) (via public-sdw-comments@w3.org)

##### ? Full use case description (click to expand):

The European INSPIRE directive leads to the availability of many datasets with geographical aspects. Unfortunately the burden for data providers to comply with INSPIRE is high, which leads to the potential danger that the means (INSPIRE) becomes more important than the end (to have interoperable data). For many data providers the sole purpose is to state that they are INSPIRE compliant. This limited but understandable approach of data providers results in potential users being overlooked. Implementations of more user-friendly Web-related solutions will be very difficult and limited in the next few years, unless we find a (documented, proven and accepted) way to comply with INSPIRE by using Web standards (Linked Data).

The Dutch Cadastre is a major provider of INSPIRE data in the Netherlands. It also maintains the portal [PDOK](#), a central facility for making INSPIRE-compliant geographical data available in the Netherlands.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.2 Compatibility with existing practices](#)

#### 4.45 Event-like geographic features

**Contributed by:** Karl Grossner, Stanford Libraries

##### ? Full use case description (click to expand):

Events are geographic phenomena. They comprise activity associated with particular locations on the earth surface, and their participants' locations and attributes over time are integral to their analysis as well as their indexing in information systems. Event participants may be human and agentive or not -including for example objects present at or resulting from activity, or natural phenomena like hurricanes or earthquakes.

In addition to class or type, and the nature of their participants, essential descriptors for events include their static or dynamic location(s) and temporal extents, described with spatial and temporal reference systems. Location may be static or dynamic. Static events routinely modeled include crime, mortality, battles, performances, etc. Temporal extents may be modeled as instants or intervals, in cases aggregated as "multi-instants" or "multi-intervals."

Although it is inherently difficult to model dynamic phenomena, there are many compelling reasons to do so. They can be conceived variously, as processes modeled as sequences of discrete events or as functions. Examples include: trajectories of people in commercial activity, journeys, battles, and biographical "life-paths."

From a spatial perspective, many places are essentially dynamic and modeling these is a core challenge in historical and cultural heritage applications. The size and shape of political entities changes over time; they grow, shrink, split, merge, disappear, reappear, etc. The point being their spatial extent is directly bound to time. Historical periods are considered as "place-at-period" - for example there is not a simple "Bronze Age" period, rather "Bronze Age Britain" and "Bronze Age Southern Levant."

Taken together these share a requirement to consider space and time together.

There is ongoing work in several research communities to arrive at better general models for representing and computing over such spatial-temporal and dynamic phenomena. One approach aims at 4-dimensional model of space-time (e.g. [CIDOC-CRM E92-Spacetime Volume](#)). [Another](#) seeks to add a "when" component to the "where" (geometry) of GeoJSON and GeoJSON-LD.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#)

**Related requirements:** [5.7 Date, time and duration](#), [5.22 4D model of space-time](#), [5.39 Space-time multi-scale](#), [5.48 Temporal reference system](#), [5.49 Temporal vagueness](#), [5.56 Valid time](#)

#### 4.46 Creation of "virtual observations" from "analysis" phase of weather prediction model

**Contributed by:** Jeremy Tandy, [Met Office](#)

##### ? Full use case description (click to expand):

National Meteorological Services (NMS), such as [Met Office](#), maintain a network of weather observation sites within their region of responsibility in order, amongst other things, to provide input to their numerical weather prediction models. The cost of maintaining these weather observation sites means that their number is limited.

Within the UK, and likely in other places too, there is a high demand for weather observations for specific locations. To meet this demand, the Met Office provides data for many more locations than the number of weather observation sites they maintain. In order to do this, "virtual observations" for these locations are derived from the "analysis" of the numerical weather prediction model. ("Analysis" is the term used to describe the initial state of the numerically modeled atmosphere from which the forecast is calculated; it incorporates real observations and provides a dynamically balanced representation of the atmosphere at a snapshot in time.)

The metadata needed to provide context to a "virtual observation" is identical to that for normal observations; albeit that the procedure used to create the observation involves a computational simulation rather than a physical sensor and stimulus. Clearly, it is important to provide information about the procedure used to create the observation so that "real" observations can be distinguished from "virtual" observations.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.3 Time Ontology in OWL](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.32 Provenance](#), [5.57 Virtual observations](#), [5.38 Sensing procedure](#), [5.52 Uncertainty in observations](#), [5.37 Sensor metadata](#), [5.6 CRS definition](#), [5.15 Georeferenced spatial data](#), [5.23 Model reuse](#), [5.35 Reference external vocabularies](#), [5.30 Observation aggregations](#), [5.51 Time series](#)

#### 4.47 Incorporating geospatial data (e.g. geo-referenced geometry) into interactive 3D graphics on the Web

**Contributed by:** Stefan Lemme, DFKI/Saarland University

##### ? Full use case description (click to expand):

This use case relates to [4.8 Consuming geographical data in a Web application](#) and maintains the bridge to 3D-web applications (without targeting the actual rendering process, which is out of scope for this Working Group).

Visualizing geospatial data, such as geo-referenced 3D geometry of buildings, is a crucial capability even for Web applications. This use case targets Web developers that are using 3D graphics in their (existing) Web applications. To ease the pick up of 3D graphics for Web developers, since they are usually non-graphics experts, the [Declarative 3D for the Web Architecture Community Group](#) of the W3C did previous work to determine the requirements, options, and use cases for integration of interactive 3D graphics capabilities into the W3C technology stack. Thereby, they propose a declarative approach to describe the 3D scene content as an extension of HTML5 rather than interfacing low-level APIs for rendering. Several implementations, such as [X3D/X3DOM](#) and [XML3D](#) address this approach. However, Web developers are usually non-geospatial experts. Thus, to achieve a similar low-entrance barrier for them to incorporate geospatial data into interactive 3D graphics on the Web, any (Web)service providing geospatial data (including geometry) might consider a compatible content delivery format and API. Firstly, this implies none to only very little processing overhead on the client. In particular, when having mobile Web applications in mind an efficient content delivery (bandwidth-wise as well as client-side-processing-wise) is becoming important. Secondly, Web developers utilize established libraries, such as jQuery, to interface remote (RESTful) APIs. To ease interfacing geospatial services, Web developers tend to reuse their tools. Finally, existing best practices of the Web should be taken into account and applied to the access of geospatial data according, such as paging of result sets, caching of resources at several stages (server, browser), etc.

1. [Examples using XML3D and OSM data \(Brussels\)](#)
2. [Examples using XML3D and OSM data \(Munich\)](#)

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related requirements:** [5.19 Linkability](#), [5.45 Streamable data](#), [5.3 Compressible](#), [5.46 Support for 3D](#), [5.50 Support for tiling](#)

#### 4.48 Smart Cities

**Contributed by:** Payam Barnaghi (on behalf of the EU FP7 CityPulse Project)

##### ? Full use case description (click to expand):

The [CityPulse project](#) has identified 101 smart city scenarios and related use-cases in cooperation with partner cities and the City Stakeholder Group in the project. The scenarios are made available online for the wider community to rank the use cases. The scenarios are [available online](#). A Large set of semantically annotated datasets collected from partners of the CityPulse EU FP7 project and



relevant resources for smart city data are [also available](#). The annotation tool for the datasets can be found [here](#).

In "101" Smart City use cases, each scenario is described in a short narrative and is ranked based on stakeholder needs and community groups by completing an online survey. The following lists top three ranked scenarios from the use cases. The complete list and details of the ranking and scenario evaluations are described in [M. Presser et al, Smart City Use-cases and Requirements](#).

1. **[Context-aware multimodal real time travel planner](#)**: Tony needs to travel from home to work. Different means of transportation are generally available to him and include walking, biking, moped/scooter (own, shared, electric/petrol), car (own, pooling, ride sharing, taxi), public transportation (bus, metro, train, ferry/boat). Transportation can be optimized to Tony's preferred travel time, convenience (comfort, seating, crowdedness, safety, environmental quality), total cost, environmental impacts, scenic route, and personal health. Factors that impact this optimization include conditions of the different transportation modes (road, weather, maintenance works, traffic intensity, people density, parking availability, charging pole availability, pollution, air quality, irregularities in traffic schedules, road tolls, seating availability, accidents, availability of city bikes,...). Tony will be presented with his ideal route and will be able to select each leg of the journey based on concurrent and projected aggregated conditions. Recalculation of his chosen route(s) can happen if conditions or preferences change, and will adapt to any detour of own choice.
2. **[Public parking space availability prediction](#)**: Frank is having a hard time finding a public parking space. The city is increasingly reducing the amount of parking spaces per unit (e.g. apartments), and the difficulty of finding a parking space means Frank has to circulate his car looking for available parking spots. This is both annoying for him and results in negative environmental impacts (pollution, noise). By using multiple input sources of information a certain degree of probability of finding a parking spot in different locations can be provided, also including timed no parking zones, thus reducing circulation times. By knowing the current situation of other circulating cars, Frank can avoid congested hot spots by being rerouted towards different paths to even out the distribution. When Frank ends up leaving his spot he can also actively provide parking availability data via his smartphone.
3. **[Stimulating green behavior](#)**: The city council has an interest in making citizens aware of their own environmental impacts in an attractive way without forcing the issue. The council provides citizens and interested parties with a dashboard, so they can follow the "environmental heartbeat" of the city as well as their own impacts and contributions. Now citizens will be able to see and measure the environmental impact on the city as a whole as well as their own personal actions, indicating noise levels, air quality, waste index, recycling activities, fresh water quality, sea water quality etc. Anyone can now compare behavior and activities to that of other groups such as neighbors, neighborhoods or city districts, resulting in competitions between these to collect "Green Points". Not only the fun of competition is stimulating, but Green Points can be traded for rebates on public transport and other environmentally friendly actions which forms a "green economy" ("green coins").

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related requirements:** [5.19 Linkability](#), [5.11 Dynamic sensor data](#), [5.16 Humans as sensors](#), [5.21 Mobile sensors](#), [5.27 Nominal observations](#), [5.18 Lightweight API](#)

#### 4.49 Provenance of climate data

**Contributed by:** Bruce Bannerman (Australian Bureau of Meteorology)

**? Full use case description (click to expand):**

This use case is inspired by one of the conclusions of the UK Parliamentary inquiry into [ClimateGate](#) "...It is not standard practice in climate science to publish the raw data and the computer code in

*academic papers. However, climate science is a matter of great importance and the quality of the science should be irreproachable. We therefore consider that climate scientists should take steps to make available all the data that support their work (including raw data) and full methodological workings (including the computer codes...)." When a climate scientist publishes a paper, he needs to be able to refer reviewers to the source data and software source code that underpins the assertions made within the paper. Climate data are typically time-series and can be quite complex. Data can be sourced from a single National Meteorological and Hydrological Service (NMHS), or from a number of NMHS. Software source code is typically stored within a software revision control repository, such as git.*

Climate data may comprise all of the following:

1. A time-series of observations of a specific phenomenon at a single sensor, including:
  - Estimations of the value of an observed property.
  - A detailed understanding of the conditions under which the observation was made.
  - 
  - Any changes that have been made to the observation (e.g. due to Quality Assurance processes).
2. A collection of the time-series observations described at #1, of the same phenomenon, at the same time steps, perhaps in the form of a discrete coverage (time-series).
3. A time-series representing the distribution of values of the collection of time-series observations represented as perhaps a:
  - Continuous two dimensional coverage.
  - Continuous coverages as three dimensional cubes.
  - Continuous coverage as n-dimensional models.
  - An ensemble, comprising a number of models.
  - The outputs of some analytical process as a time-series/coverage/cube/model.

So using the description of climate data above, when a paper is published, the scientist needs to be able to refer viewers to:

- The analytical data which underpin the paper (perhaps the n-dimensional continuous coverage time-series at #3).
- The quality assured observations data that the continuous coverages at #3 were derived from at #2 and #1 (with details as to why each change to the Quality Assured observations described at #1 were made).
- The 'raw' observations data described at #1, with details as to the conditions, sensors, etc., that the observation was made under.
- The version of the software source code that was used to manipulate the data at #1, #2, and #3.

This is not a trivial data management problem to address, however its resolution will provide a solid data management grounding for future climate science and help address much spurious debate. Parts of the puzzle are currently being worked on, e.g.:

- The World Meteorological Organisation (WMO) has published [WMO No. 1131: Climate Data Management Systems Specifications](#) that provide a high level conceptual architecture to address much of the data management issues described above.
- WMO and OGC have developed and are developing relevant logical data models ISO 19156 Observations and Measurements, OGC Timeseries, and WMO METCE. The last two are based on Observations and Measurements.
- WMO has released a standard for describing Observations Metadata, called WIGOS Metadata.
- WMO is about to start work on a logical data model based on Observations and Measurements and Timeseries for describing WMO Observations. A future iteration of this model will need to cater for data provenance.
- [W3C](#) has developed PROV-O, which has considerable potential for describing data provenance.

The missing part is: How can the provenance of the collection of climate data and the software used to manipulate it be best modeled and in the future, found via the Internet? The resolution of this issue will be of relevance to many domains.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.10 Discoverability](#), [5.23 Model reuse](#), [5.30 Observation aggregations](#), [5.31 Observed property in coverage](#), [5.32 Provenance](#), [5.33 Quality metadata](#), [5.37 Sensor metadata](#), [5.38 Sensing procedure](#), [5.40 Spatial metadata](#), [5.46 Support for 3D](#), [5.51 Time series](#), [5.57 Virtual observations](#)

## 4.50 Representing geospatial data in RDF

**Contributed by:** Phil Archer on behalf of the SmartOpendata project

### ? Full use case description (click to expand):

Beginning in late 2013, the EU-funded [SmartOpenData project](#) carried out a number of pilots examining the potential of Linked Data for environmental and rural issues, particularly in National Parks. One aspect of the project required parts of the European Union's [INSPIRE data model](#) to be represented in RDF and used across the different pilots.

The [INSPIRE](#) data model is large, complex, and designed for use in a Geospatial Information System and not for Linked Data. Initially, the model was followed directly, creating classes and properties that exactly matched those in the original [INSPIRE](#) model. However, this produced cumbersome RDF that was hard to use, was difficult to link to other data points, and offered no advantage over the original. It was using Linked Data for the sake of it.

As a result, the decision was taken not to attempt to replicate the whole model in RDF but to follow a more Linked Data centric approach, re-using concepts wherever necessary but simplifying it where possible, and making maximum use of external URI sets. An important reference in this work was the [Study on RDF and PIDs for INSPIRE](#) by Diederik Tirry and Danny Vandenbroucke. It summarized work by three experts: Clemens Portele, Linda van den Brink and Stuart Williams. The establishment of the [INSPIRE Registry](#), that provides stable URIs for many of the code lists was also a key development.

As a result of the work by the European Commission and the SmartOpenData project, a [set of small RDF vocabularies](#) is available that represents a subset of the [INSPIRE](#) model. The desire to see these extended by future projects is stated explicitly.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#) [2.5 Coverage in Linked Data](#)

**Related requirements:** [5.2 Compatibility with existing practices](#), [5.19 Linkability](#), [5.23 Model reuse](#), [5.25 Multilingual support](#), [5.35 Reference external vocabularies](#)

## 5. Accepted requirements

This chapter lists the requirements for the deliverables of the Working Group, in alphabetical order.

NOTE

In some requirements the expression 'recommended way' is used. This means that a single best way of doing something is sought. It does not say anything about the form this recommended way should have, or who should make the recommendation. A recommended way could be a formal recommendation or standard from an authoritative standards body like the OGC or W3C, but it could just as well be a more informal specification, as long as it is arguably the best way of doing something.

## 5.1 Bounding box and centroid

There should be a recommended way for publishing and requesting the bounding box and centroid of a spatial thing.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.7 Publishing geographical data](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#)

## 5.2 Compatibility with existing practices

Standards or recommendations for spatial data on the Web should be compatible with existing methods of making spatial data available (like WFS, WMS, CSW, WCS).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.7 Publishing geographical data](#), [4.12 Using spatial data from the Web in GIS systems during emergency response operations](#), [4.44 INSPIRE compliance using Web standards](#), [4.50 Representing geospatial data in RDF](#)

## 5.3 Compressible

Spatial data on the Web should be compressible (for optimization of data transfer).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.16 Dutch Base Registry](#), [4.21 Driving to work in the snow](#), [4.47 Incorporating geospatial data \(e.g. geo-referenced geometry\) into interactive 3D graphics on the Web](#)

## 5.4 Coverage temporal extent

It should be possible to add temporal references to spatial coverage data.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

[4.19 Publication of raw subsurface monitoring data](#), [4.3 Real-time wildfire monitoring](#), [4.32 Satellite data processing](#), [4.30 Spatial sampling](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#),

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.1 Meteorological data rescue](#), [4.17 Publishing Cultural heritage data](#), [4.40 TCGA / microscopy imaging](#), [4.37 Landsat data services](#), [4.21 Driving to work in the snow](#)

## 5.5 Crawlability

Spatial data on the Web should be crawlable, allowing data to be found and indexed by external agents.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.1 Meteorological data rescue](#), [4.5 Harvesting of Local Search Content](#), [4.8 Consuming geographical data in a Web application](#), [4.39 Crowdsourced earthquake observation information](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#)

## 5.6 CRS definition

There should be a recommended way of referencing a Coordinate Reference System (CRS) with a HTTP URI, and to get useful information about the CRS when that URI is dereferenced.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.4 Diachronic burnt scar Mapping](#), [4.1 Meteorological data rescue](#), [4.25 Images, e.g. a time series of a water course](#), [4.41 Crop yield estimation using multiple satellites](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#)

## 5.7 Date, time and duration

It should be possible to represent dates, time and duration.

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.21 Driving to work in the snow](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.45 Event-like geographic features](#), [4.5 Harvesting of Local Search Content](#), [4.22 Intelligent Transportation System](#), [4.37 Landsat data services](#), [4.1 Meteorological data rescue](#), [4.29 Observations on geological samples](#), [4.13 Publication of air quality data aggregations](#), [4.31 Select hierarchical geographical regions for use in data analysis or visualisation](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.17 Publishing Cultural heritage data](#), [4.25 Images, e.g. a time series of a water course](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#)

## 5.8 Default CRS

There should be a default Coordinate Reference System (CRS) that can be assumed to be used for coordinates that otherwise have no specification of the CRS.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.15 Combining spatial RDF data for integrated querying in a triplestore](#)

### ISSUE 28

This requirement is under discussion.

## 5.9 Different time models

It should be possible to represent data using different time models, such as geological time and non-Gregorian calendars.

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.38 Metadata and search granularity](#), [4.1 Meteorological data rescue](#), [4.29 Observations on geological samples](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#)

## 5.10 Discoverability

It should be easy to find spatial data on the Web, e.g. by means of metadata aimed at discovery. When spatial data are published on the Web, both humans and machines should be able to discover those data.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.28 Bushfire response coordination centre](#), [4.8 Consuming geographical data in a Web application](#), [4.41 Crop yield estimation using multiple satellites](#), [4.18 Dissemination of 3D geological data](#), [4.21 Driving to work in the snow](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.5 Harvesting of Local Search Content](#), [4.25 Images, e.g. a time series of a water course](#), [4.43 Improving discovery of spatial data on the Web](#), [4.11 Integration of governmental and utility data to enable smart grids](#), [4.22 Intelligent Transportation System](#), [4.37 Landsat data services](#), [4.33 Marine observations - eMII](#), [4.1 Meteorological data rescue](#), [4.38 Metadata and search granularity](#), [4.27 Soil data applications](#), [4.12 Using spatial data from the Web in GIS systems during emergency response operations](#), [4.40 TCGA / microscopy imaging](#), [4.39 Crowdsourced earthquake observation information](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.49 Provenance of climate data](#)

## 5.11 Dynamic sensor data

It should be possible to represent near real-time streaming sensor measurements.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.39 Crowdsourced earthquake observation information](#), [4.3 Real-time wildfire monitoring](#), [4.21 Driving to work in the snow](#), [4.26 Droughts in geological complex environments where groundwater is important](#), [4.19 Publication of raw subsurface monitoring data](#), [4.48 Smart Cities](#)

## 5.12 Encoding for vector geometry

There should be a recommended way of encoding vector geometry (an expression of spatial data that uses coordinates) when such data are published on the Web.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.7 Publishing geographical data](#), [4.8 Consuming geographical data in a Web application](#), [4.15 Combining spatial RDF data for integrated querying in a triplestore](#), [4.22 Intelligent Transportation System](#), [4.44 INSPIRE compliance using Web standards](#), [4.17 Publishing Cultural heritage data](#), [4.39 Crowdsourced earthquake observation information](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#)



### 5.13 Ex-situ sampling

It should be possible to represent ex-situ (remote) sampling or sensing.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.38 Metadata and search granularity](#), [4.29 Observations on geological samples](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#)

### 5.14 Georectification

The coverage data model should consider the inclusion of metadata to allow georectification to an arbitrary grid.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.4 Diachronic burnt scar Mapping](#), [4.3 Real-time wildfire monitoring](#)

### 5.15 Georeferenced spatial data

It should be possible to georeference spatial data.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.18 Dissemination of 3D geological data](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.25 Images, e.g. a time series of a water course](#), [4.1 Meteorological data rescue](#), [4.17 Publishing Cultural heritage data](#), [4.32 Satellite data processing](#), [4.41 Crop yield estimation using multiple satellites](#), [4.39 Crowdsourced earthquake observation information](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.21 Driving to work in the snow](#), [4.25 Images, e.g. a time series of a water course](#), [4.29 Observations on geological samples](#), [4.30 Spatial sampling](#), [4.32 Satellite data processing](#), [4.34 Marine observations - data providers](#), [4.19 Publication of raw subsurface monitoring data](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.10 Publishing geospatial reference data](#)

### 5.16 Humans as sensors

It should be possible to represent observations taken by human individuals or communities acting as sensors perceiving the environment.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.21 Driving to work in the snow](#), [4.32 Satellite data processing](#), [4.34 Marine observations - data providers](#), [4.39 Crowdsourced earthquake observation information](#), [4.19 Publication of raw subsurface monitoring data](#), [4.29 Observations on geological samples](#), [4.27 Soil data applications](#), [4.48 Smart Cities](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#)

### 5.17 Independence on reference systems

Standards or recommendations for spatial data on the Web should be independent on the reference systems that are used for data.

*Note: This requirement reflects that spatial data incorporate geographical data, but can also be data that are not directly related to earth or its scale.*

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.6 Locating a thing](#), [4.8 Consuming geographical data in a Web application](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.32 Satellite data processing](#), [4.40 TCGA / microscopy imaging](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#)

## 5.18 Lightweight API

A lightweight API is needed for implementation on Internet of Things (IoT) devices.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.21 Driving to work in the snow](#), [4.48 Smart Cities](#)

### ISSUE 24

This requirement needs clarification.

## 5.19 Linkability

Spatial data on the Web should be linkable (by explicit relationships between different data in different data sets), to other spatial data and to or from other types of data.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.28 Bushfire response coordination centre](#), [4.41 Crop yield estimation using multiple satellites](#), [4.4 Diachronic burnt scar Mapping](#), [4.26 Droughts in geological complex environments where groundwater is important](#), [4.29 Observations on geological samples](#), [4.23 Optimizing energy consumption, production, sales and purchases in Smart Grids](#), [4.3 Real-time wildfire monitoring](#), [4.30 Spatial sampling](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.10 Publishing geospatial reference data](#), [4.16 Dutch Base Registry](#), [4.1 Meteorological data rescue](#), [4.19 Publication of raw subsurface monitoring data](#), [4.48 Smart Cities](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.25 Images, e.g. a time series of a water course](#), [4.39 Crowdsourced earthquake observation information](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#), [4.50 Representing geospatial data in RDF](#), [4.47 Incorporating geospatial data \(e.g. geo-referenced geometry\) into interactive 3D graphics on the Web](#)

## 5.20 Machine to machine

Standards or recommendations for spatial data on the Web should work well in machine to machine environments.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.6 Locating a thing](#), [4.21 Driving to work in the snow](#), [4.39 Crowdsourced earthquake observation information](#), [4.43 Improving discovery of spatial data on the Web](#)

## 5.21 Mobile sensors

It should be possible to represent sensors that change their location, as well as the current location of the sensor at the observation time.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.39 Crowdsourced earthquake observation information](#), [4.21 Driving to work in the snow](#), [4.22 Intelligent Transportation System](#), [4.33 Marine observations - eMII](#), [4.13 Publication of air quality data aggregations](#), [4.30 Spatial sampling](#), [4.32 Satellite data processing](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.48 Smart Cities](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#)

## 5.22 4D model of space-time

It should be possible to represent spatial extent directly bound to time, e.g. journey trajectories.

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.45 Event-like geographic features](#), [4.19 Publication of raw subsurface monitoring data](#)

## 5.23 Model reuse

Spatial data modeling issues solved in existing models shall be considered for adoption, e.g. O&M, SoilML or the OGC coverage model.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.33 Marine observations - eMII](#), [4.18 Dissemination of 3D geological data](#), [4.29 Observations on geological samples](#), [4.13 Publication of air quality data aggregations](#), [4.27 Soil data applications](#), [4.30 Spatial sampling](#), [4.19 Publication of raw subsurface monitoring data](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.49 Provenance of climate data](#), [4.50 Representing geospatial data in RDF](#)

### ISSUE 18

This requirement needs clarification.

## 5.24 Moving features

It should be possible to refer to features that change their location.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.21 Driving to work in the snow](#), [4.22 Intelligent Transportation System](#)

## 5.25 Multilingual support

All vocabularies that will be developed or revised should have annotation in multiple languages.

### NOTE

We assume that it is technically possible to have human readable annotation in multiple languages in vocabularies for spatiotemporal data on the Web. With English being the default language, as many other languages as possible should be supported. This will mean that in such cases we will have to call upon WG members that are fluent in languages other than English to provide annotation.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#), [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.1 Meteorological data rescue](#), [4.39 Crowdsourced earthquake observation information](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#), [4.50 Representing geospatial data in RDF](#)

## 5.26 Multiple types of coverage

It should be possible to represent many different types of coverage. For instance, to classify coverage data by grid complexity: GridCoverage (GML 3.2.1), RectifiedGridCoverage, ReferenceableGridCoverage, etc.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.4 Diachronic burnt scar Mapping](#), [4.18 Dissemination of 3D geological data](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.25 Images, e.g. a time series of a water course](#), [4.3 Real-time wildfire monitoring](#), [4.32 Satellite data processing](#), [4.27 Soil data applications](#)

### ISSUE 19

This requirement needs clarification.

## 5.27 Nominal observations

it should be possible to represent qualitative and nominal observations.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.39 Crowdsourced earthquake observation information](#), [4.19 Publication of raw subsurface monitoring data](#), [4.32 Satellite data processing](#), [4.27 Soil data applications](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.48 Smart Cities](#)

## 5.28 Nominal temporal references

It should be possible to refer to time intervals by nominal temporal references (e.g. January, a named event in a calendar, a geological period, a dynastic period).

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.21 Driving to work in the snow](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.38 Metadata and search granularity](#), [4.1 Meteorological data rescue](#), [4.17 Publishing Cultural heritage data](#), [4.20 Use of a place name](#)

[ontology for geo-parsing text and geo-enabling searches](#), [4.43 Improving discovery of spatial data on the Web](#)

## 5.29 Non-geographic reference system

It should be possible to use non-geographic spatial reference systems.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.17 Publishing Cultural heritage data](#), [4.40 TCGA / microscopy imaging](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#)

## 5.30 Observation aggregations

It should be possible to represent aggregations of observations.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.39 Crowdsourced earthquake observation information](#), [4.33 Marine observations - eMII](#), [4.13 Publication of air quality data aggregations](#), [4.19 Publication of raw subsurface monitoring data](#), [4.27 Soil data applications](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.49 Provenance of climate data](#)

## 5.31 Observed property in coverage

It should be possible to describe the observed property represented by a coverage.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.18 Dissemination of 3D geological data](#), [4.33 Marine observations - eMII](#), [4.1 Meteorological data rescue](#), [4.27 Soil data applications](#), [4.49 Provenance of climate data](#)

## 5.32 Provenance

Ensure alignment of models or vocabularies for describing provenance that exist in the geospatial and Semantic Web domains. Examples are the [W3C](#) provenance ontology (PROV-O) and the OGC metadata specification (ISO-19115).

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#), [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.41 Crop yield estimation using multiple satellites](#), [4.4 Diachronic burnt scar Mapping](#), [4.26 Droughts in geological complex environments where groundwater is important](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.1 Meteorological data rescue](#), [4.29 Observations on geological samples](#), [4.3 Real-time wildfire monitoring](#), [4.32 Satellite data processing](#), [4.25 Images, e.g. a time series of a water course](#), [4.40 TCGA / microscopy imaging](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#), [4.49 Provenance of climate data](#)

## 5.33 Quality metadata

It should be possible to describe properties of the data quality, e.g. uncertainty.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.1 Meteorological data rescue](#), [4.18 Dissemination of 3D geological data](#), [4.26 Droughts in geological complex environments where groundwater is important](#), [4.33 Marine observations - eMII](#), [4.27 Soil data applications](#), [4.25 Images, e.g. a time series of a water course](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#), [4.49 Provenance of climate data](#)

### 5.34 Reference data chunks

It should be possible to identify and reference chunks of data, e.g. for processing, citation, provenance, cataloging.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.18 Dissemination of 3D geological data](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.25 Images, e.g. a time series of a water course](#), [4.37 Landsat data services](#), [4.1 Meteorological data rescue](#), [4.40 TCGA / microscopy imaging](#)

### 5.35 Reference external vocabularies

It should be possible to refer to externally-managed controlled vocabularies.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.13 Publication of air quality data aggregations](#), [4.19 Publication of raw subsurface monitoring data](#), [4.29 Observations on geological samples](#), [4.27 Soil data applications](#), [4.1 Meteorological data rescue](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.25 Images, e.g. a time series of a water course](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.50 Representing geospatial data in RDF](#)

#### ISSUE 20

This requirement could be out of scope.

### 5.36 Sampling topology

It should be possible to represent topological relationships between observation samples, e.g. specimens located along a borehole or probe spots found on a polished section of rocks.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.30 Spatial sampling](#), [4.33 Marine observations - eMII](#), [4.19 Publication of raw subsurface monitoring data](#), [4.29 Observations on geological samples](#), [4.25 Images, e.g. a time series of a water course](#)

### 5.37 Sensor metadata

It should be possible to include metadata about the sensors producing the observations.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)



**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.39 Crowdsourced earthquake observation information](#), [4.1 Meteorological data rescue](#), [4.3 Real-time wildfire monitoring](#), [4.4 Diachronic burnt scar Mapping](#), [4.11 Integration of governmental and utility data to enable smart grids](#), [4.21 Driving to work in the snow](#), [4.28 Bushfire response coordination centre](#), [4.29 Observations on geological samples](#), [4.30 Spatial sampling](#), [4.32 Satellite data processing](#), [4.33 Marine observations - eMII](#), [4.19 Publication of raw subsurface monitoring data](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.25 Images, e.g. a time series of a water course](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.49 Provenance of climate data](#)

### 5.38 Sensing procedure

It should be possible to attach the procedural description of a sensing method.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.32 Satellite data processing](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.1 Meteorological data rescue](#), [4.19 Publication of raw subsurface monitoring data](#), [4.35 Marine observations - data consumers](#), [4.29 Observations on geological samples](#), [4.27 Soil data applications](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.49 Provenance of climate data](#)

### 5.39 Space-time multi-scale

It should be possible to represent and integrate data over spatial and temporal scales.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.25 Images, e.g. a time series of a water course](#), [4.22 Intelligent Transportation System](#), [4.1 Meteorological data rescue](#), [4.27 Soil data applications](#), [4.30 Spatial sampling](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.45 Event-like geographic features](#), [4.38 Metadata and search granularity](#), [4.17 Publishing Cultural heritage data](#), [4.19 Publication of raw subsurface monitoring data](#), [4.49 Provenance of climate data](#)

### 5.40 Spatial metadata

There should be recommended ways for describing the spatial characteristics of data (data objects, data sets or data services), like:

- number of dimensions (1D, 2D, 3D)
- data type (raster or vector)
- Coordinate Reference System(s)
- spatial extent
- spatial resolution
- positioning system used for acquiring location

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.1 Meteorological data rescue](#), [4.5 Harvesting of Local Search Content](#), [4.7 Publishing geographical data](#), [4.8 Consuming geographical data in a Web application](#), [4.12 Using spatial data from the Web in GIS systems during emergency response operations](#), [4.15 Combining spatial RDF data for integrated querying in a triplestore](#), [4.19 Publication of raw subsurface monitoring data](#), [4.22 Intelligent Transportation System](#), [4.25 Images, e.g. a time series of a water course](#), [4.27 Soil data applications](#), [4.37 Landsat data services](#), [4.38 Metadata and search granularity](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving](#)

[discovery of spatial data on the Web](#), [4.44 INSPIRE compliance using Web standards](#), [4.40 TCGA / microscopy imaging](#), [4.18 Dissemination of 3D geological data](#), [4.39 Crowdsourced earthquake observation information](#)

### 5.41 Spatial relationships

There should be a recommended way for expressing spatial relationships between spatial entities. These relationships can be topological, mereological (part-whole), directional or distance related.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.6 Locating a thing](#), [4.7 Publishing geographical data](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.31 Select hierarchical geographical regions for use in data analysis or visualisation](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.36 Building information management and data sharing](#)

### 5.42 Spatial operators

There should be a recommended way for the definition and use of spatial operators. Spatial things can have spatial relations: topological relations, directional or distance relations. Operators based on these relations (e.g. 'Contains', 'Intersects', 'Nearest') should be well defined and easy to use.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.8 Consuming geographical data in a Web application](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.31 Select hierarchical geographical regions for use in data analysis or visualisation](#), [4.38 Metadata and search granularity](#), [4.43 Improving discovery of spatial data on the Web](#), [4.10 Publishing geospatial reference data](#), [4.40 TCGA / microscopy imaging](#), [4.41 Crop yield estimation using multiple satellites](#), [4.39 Crowdsourced earthquake observation information](#)

#### ISSUE 23

This requirement needs clarification.

### 5.43 Spatial vagueness

It should be possible to describe locations in a vague, imprecise manner. For instance, to represent spatial descriptions from crowdsourced observations, such as "we saw a wildfire at the bottom of the hillside" or "we felt a light tremor while walking by Los Angeles downtown". Another related use case deals with spatial locations identified in historical texts, e.g. a battle occurred at the south west boundary of the Roman Empire.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.28 Bushfire response coordination centre](#), [4.38 Metadata and search granularity](#), [4.1 Meteorological data rescue](#), [4.30 Spatial sampling](#), [4.39 Crowdsourced earthquake observation information](#), [4.17 Publishing Cultural heritage data](#), [4.19 Publication of raw subsurface monitoring data](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#)

## 5.44 SSN-like representation

It should be possible to represent satellite data using the SSN model, including sensor descriptions.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.4 Diachronic burnt scar Mapping](#), [4.3 Real-time wildfire monitoring](#)

## 5.45 Streamable data

Data should be streamable, a consumer should be able to do something meaningful before the end of the data message is received. This could be considered a general requirement for data on the Web, but it is recorded here because spatial data often consist of large chunks of data.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.8 Consuming geographical data in a Web application](#), [4.47 Incorporating geospatial data \(e.g. geo-referenced geometry\) into interactive 3D graphics on the Web](#)

## 5.46 Support for 3D

Standards or recommendations for spatial data on the Web should be applicable to three-dimensional data.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#), [2.4 Semantic Sensor Network Vocabulary](#), [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.18 Dissemination of 3D geological data](#), [4.7 Publishing geographical data](#), [4.19 Publication of raw subsurface monitoring data](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.25 Images, e.g. a time series of a water course](#), [4.26 Droughts in geological complex environments where groundwater is important](#), [4.36 Building information management and data sharing](#), [4.39 Crowdsourced earthquake observation information](#), [4.40 TCGA / microscopy imaging](#), [4.33 Marine observations - eMII](#), [4.34 Marine observations - data providers](#), [4.41 Crop yield estimation using multiple satellites](#), [4.38 Metadata and search granularity](#), [4.2 Habitat zone verification for designation of Marine Conservation Zones](#), [4.49 Provenance of climate data](#), [4.47 Incorporating geospatial data \(e.g. geo-referenced geometry\) into interactive 3D graphics on the Web](#)

## 5.47 Time dependencies in CRS definitions

It should be possible for Coordinate Reference Systems to have time dependent components such as the point of origin.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.6 Locating a thing](#), [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.14 Publication of transport card validation and recharging data](#)

## 5.48 Temporal reference system

If a temporal reference is used, the definition of the temporal reference system (e.g. Unix date, Gregorian Calendar, Japanese Imperial Calendar, Carbon Date, Geological Date) should be referenceable online.

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.45 Event-like geographic features](#), [4.22 Intelligent Transportation System](#), [4.38 Metadata and search granularity](#), [4.1 Meteorological data rescue](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.29 Observations on geological samples](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#)

### 5.49 Temporal vagueness

It should be possible to describe time points and intervals in a vague, imprecise manner. For instance, to represent an event happened on the afternoon of June 1st or at the second quarter of the 9th century.

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.9 Enabling publication, discovery and analysis of spatiotemporal data in the humanities](#), [4.45 Event-like geographic features](#), [4.17 Publishing Cultural heritage data](#), [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.43 Improving discovery of spatial data on the Web](#)

#### ISSUE 26

This requirement needs clarification.

### 5.50 Support for tiling

Standards or recommendations for spatial data on the Web should support tiling (for raster and vector data). Tiling of spatial data can drastically improve the speed of data retrieval and allows having simple caches of data around the Web.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.8 Consuming geographical data in a Web application](#), [4.47 Incorporating geospatial data \(e.g. geo-referenced geometry\) into interactive 3D graphics on the Web](#)

### 5.51 Time series

It should be possible to represent time series of data.

**Related deliverables:** [2.3 Time Ontology in OWL](#), [2.5 Coverage in Linked Data](#), [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.26 Droughts in geological complex environments where groundwater is important](#), [4.25 Images, e.g. a time series of a water course](#), [4.37 Landsat data services](#), [4.19 Publication of raw subsurface monitoring data](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.25 Images, e.g. a time series of a water course](#), [4.40 TCGA / microscopy imaging](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#), [4.49 Provenance of climate data](#)

### 5.52 Uncertainty in observations

It should be possible to represent uncertainty in observations.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.34 Marine observations - data providers](#), [4.1 Meteorological data rescue](#), [4.19 Publication of raw subsurface monitoring data](#), [4.30 Spatial sampling](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#)

### 5.53 Update datatypes in OWL Time

OWL Time should be updated to align with the 2012 update of OWL datatypes and 2012 update of xsd datatypes

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.17 Publishing Cultural heritage data](#)

### 5.54 Use in computational models

It should be possible to use coverage data as input or output of computational models, e.g. geological models.

**Related deliverables:** [2.5 Coverage in Linked Data](#)

**Related use cases:** [4.41 Crop yield estimation using multiple satellites](#), [4.18 Dissemination of 3D geological data](#), [4.26 Droughts in geological complex environments where groundwater is important](#)

### 5.55 Validation

It should be possible to validate spatial data on the Web; to automatically detect conflicts with standards or definitions.

**Related deliverables:** [2.2 Spatial Data on the Web Best Practices](#)

**Related use cases:** [4.10 Publishing geospatial reference data](#), [4.17 Publishing Cultural heritage data](#), [4.36 Building information management and data sharing](#)

### 5.56 Valid time

Ensure alignment with existing methods for expressing the time in which data are valid (e.g. dcterms:valid)

**Related deliverables:** [2.3 Time Ontology in OWL](#)

**Related use cases:** [4.20 Use of a place name ontology for geo-parsing text and geo-enabling searches](#), [4.25 Images, e.g. a time series of a water course](#), [4.31 Select hierarchical geographical regions for use in data analysis or visualisation](#), [4.45 Event-like geographic features](#), [4.42 Geospatial extensions to domain-independent metadata schemas](#)

#### NOTE

This requirement does not mean that the way expressions of time can be used is considered in scope for the Time Ontology, but it does mean that being able to express temporal validity using the Time Ontology is considered important for location data.

## 5.57 Virtual observations

It must be possible to represent synthetic observations made by computational procedures or inference.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.26 Droughts in geological complex environments where groundwater is important](#), [4.27 Soil data applications](#), [4.32 Satellite data processing](#), [4.41 Crop yield estimation using multiple satellites](#), [4.46 Creation of “virtual observations” from “analysis” phase of weather prediction model](#), [4.19 Publication of raw subsurface monitoring data](#), [4.49 Provenance of climate data](#)

## 6. Accepted requirements by deliverable

For convenience, this chapter lists requirements grouped by Working Group deliverable.

### 6.1 Requirements for deliverable [Spatial Data on the Web Best Practices](#):

1. [Bounding box and centroid](#)
2. [Compatibility with existing practices](#)
3. [Compressible](#)
4. [Crawlability](#)
5. [CRS definition](#)
6. [Default CRS](#)
7. [Discoverability](#)
8. [Encoding for vector geometry](#)
9. [Independence on reference systems](#)
10. [Linkability](#)
11. [Machine to machine](#)
12. [Multilingual support](#)
13. [Provenance](#)
14. [Spatial metadata](#)
15. [Spatial relationships](#)
16. [Spatial operators](#)
17. [Streamable data](#)
18. [Support for 3D](#)
19. [Time dependencies in CRS definitions](#)
20. [Support for tiling](#)
21. [Validation](#)

### 6.2 Requirements for deliverable [Time Ontology in OWL](#):

1. [Date, time and duration](#)
2. [Different time models](#)
3. [4D model of space-time](#)
4. [Multilingual support](#)
5. [Nominal temporal references](#)
6. [Provenance](#)
7. [Space-time multi-scale](#)
8. [Temporal reference system](#)
9. [Temporal vagueness](#)
10. [Time series](#)
11. [Update datatypes in OWL Time](#)
12. [Valid time](#)



### 6.3 Requirements for deliverable [Semantic Sensor Network Vocabulary](#):

1. [Compatibility with existing practices](#)
2. [CRS definition](#)
3. [Dynamic sensor data](#)
4. [Ex-situ sampling](#)
5. [Georeferenced spatial data](#)
6. [Humans as sensors](#)
7. [Independence on reference systems](#)
8. [Lightweight API](#)
9. [Linkability](#)
10. [Machine to machine](#)
11. [Mobile sensors](#)
12. [Model reuse](#)
13. [Moving features](#)
14. [Multilingual support](#)
15. [Nominal observations](#)
16. [Observation aggregations](#)
17. [Provenance](#)
18. [Reference external vocabularies](#)
19. [Sampling topology](#)
20. [Sensor metadata](#)
21. [Sensing procedure](#)
22. [Space-time multi-scale](#)
23. [Spatial vagueness](#)
24. [Support for 3D](#)
25. [Time series](#)
26. [Uncertainty in observations](#)
27. [Virtual observations](#)

### 6.4 Requirements for deliverable [Coverage in Linked Data](#):

1. [Coverage temporal extent](#)
2. [Discoverability](#)
3. [Georectification](#)
4. [Georeferenced spatial data](#)
5. [Model reuse](#)
6. [Multilingual support](#)
7. [Multiple types of coverage](#)
8. [Non-geographic reference system](#)
9. [Observed property in coverage](#)
10. [Provenance](#)
11. [Quality metadata](#)
12. [Reference data chunks](#)
13. [Spatial vagueness](#)
14. [SSN-like representation](#)
15. [Support for 3D](#)
16. [Time series](#)
17. [Use in computational models](#)

## 7. Deferred requirements

This section lists requirements that are considered relevant for the group's work, but will not be addressed in the current time frame.

## 7.1 Definition of SSN profiles and checking compliance against the model

There should be a recommended way for defining profiles of the SSN model (i.e. "a constrained subset of a more general model", see [glossary](#)) and for checking compliance of profiles against the SSN model.

**Related deliverables:** [2.4 Semantic Sensor Network Vocabulary](#)

**Related use cases:** [4.32 Satellite data processing](#)

### [ISSUE 13](#)

Discussion related to this requirement can be found at the issue page.

## 8. Acknowledgements

The editors are grateful for all contributions made to this document, in particular the use case authors.

Copyright © 2015 by [DFKI](#) for examples in [Incorporating geospatial data into interactive 3D graphics on the Web](#). The [MIT License](#) applies.

## 9. Appendix A: History of changes

### 9.1 Changes between the First Public Working Draft and Second Public Working Draft

A summary of the main changes between the [First Public Working Draft](#) (published 2015-07-23) and the Second Public Working Draft (this version) of this document:

- A note was added to the description of the deliverable [Coverage In Linked Data](#).
- Two new use cases were added: [Provenance of climate data](#) and [Representing geospatial data in RDF](#).
- A note was added to the [section listing accepted requirements](#), to explain the term 'recommended way'.
- The [Provenance requirement](#) was rephrased.
- The [Valid time requirement](#) was rephrased.
- The requirement 'Spatial meronymy' was renamed to [Spatial relationships](#) and redefined.
- A new requirement for the deliverable [Time Ontology in OWL](#) was added: [Update datatypes in OWL Time](#).
- A new section was added that lists [accepted requirements by deliverable](#).
- A new section listing [deferred requirements](#) was added.