

ESIP Summer Meeting 2019 Advancing spatial and temporal aspects of schema.org

Charles Vardeman, Adam Shephard and Doug Fills *cvardema@nd.edu*







Funded by the National Science Foundation Grant #1842042

Schema.org Neon Vocamp's

Charles Vardeman and Jane Wyngaard, UND

Christine Laney and Tom Gulbransen, NEON/Battelle

Adam Shepherd, BCO-DMO/Woods Hole Oceanographic Institution

Doug Flls, Consortium for Ocean Leadership







CENTER FOR APPLIED CYBERSECURITY RESEARCH

ψ







Funded by the National Science Foundation Grant #1842042

Cyberinfrastructure Center of Excellence Pilot

Ewa Deelman, USC (PI)

Co-PIs: Anirban Mandal, RENCI Jarek Nabrzyski, Notre Dame University Valerio Pascucci and Rob Ricci, University of Utah







CENTER FOR APPLIED CYBERSECURITY RESEARCH

Ш



PROBLEM: schema.org is currently inconsistent with standards organizations (W3C, OGC) representations of spatial and temporal information. This session will bring together data curators, conceptual modelers and ontologists to formulate solutions for extending schema.org's approach to spatial and temporal descriptions.





Thrust:

Team members: Anirban Mandal, Laura Christopherson, Erik Scott, Ilya Baldin, Paul Ruth













- Interactive vocamp unconference style session
- schema.org place
- connection to existing standards
- schema.org spatial definitions

Bioschemas.org as a model?





🗥 Home Getting Started Specifications 🔻 Deploy & Develop ୟ

What is Bioschemas?

Bioschemas aims to improve the Findability of data in the life sciences. It does this by encouraging people in the life sciences to use Schema.org markup in their websites so that they are indexable by search engines and other services. Bioschemas encourages the consistent use of markup to ease the consumption of the contained markup across many sites. This structured information then makes it easier to discover, collate, and analyse distributed data.

Bioschemas is making two main contributions:

- Proposing new types and properties to Schema.org to allow for the description of life science resources.
- 2. Profiles over the Schema.org types that identify the essential properties to use in describing a resource.

Bioschemas started as a community effort in November 2015. It operates as an open community initiative with representatives from a wide variety of institutions. You are welcome to join the community.

Schema.org

Schema.org is a community effort supported by the main search engines, and is already widely implemented across the web.

Bioschemas.org Profiles



Name	Group	Use Cases	Cross Walk	Task & Issues	Examples	Live Deploys
DataCatalog (v0.3) 01 July 2019	Data Repositories	Ņ	×			۶
<mark>Dataset</mark> (v0.3) 14 June 2019	Datasets	Ļ	Ķ			۶
Event (v0.1) 05 July 2018	Events	Ļ	Ķ			۶
<mark>Sample</mark> (v0.2) 10 November 2018	Samples		Ķ			۶
<mark>Taxon</mark> (v0.3) 10 November 2018	Biodiversity		×.			F
<mark>Tool</mark> (v0.1) 07 March 2018	Tools	Ę.	Ŕ			۶
TrainingMaterial (v0.2) 05 July 2018	Training		×.			F

Bioschemas "Type Extensions"



Name	Group	Task & Issues
BioChemEntity (v0.7-RC) 2019-06-14	Biological Entities	
BioSample (v0.1-RC) 2019-06-14	Samples	
ChemicalSubstance (v0.2-RC) 2019-06-14	Chemicals	
Gene (v0.2-RC) 2019-06-14	Genes	
MolecularEntity (v0.2-RC) 2019-06-14	Chemicals	
Protein (v0.2-RC) 2019-06-14	Proteins	
Taxon (v0.3-RC) 2019-06-14	Biodiversity	



What would be the equivalent "Types" for geoschemas.org that are not covered by schema.org?

schema.org/Place



schema.org	Custom Search
	Home Schemas Documentation
Diaco	

Place

Thing > Place

Entities that have a somewhat fixed, physical extension.

[more...]

Property	Expected Type	Description
Properties from Place		
	PropertyValue	A property-value pair representing an additional characteristics of the entitity, e.g. a product feature or another characteristic for which there is no matching property in schema.org.
additionalProperty		Note: Publishers should be aware that applications designed to use specific schema.org properties (e.g. http://schema.org/width, http://schema.org/color, http://schema.org/gtin13,) will typically expect such data to be provided using those properties, rather than using the generic property/value mechanism.
address	PostalAddress or Text	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
	LocationFeatureSpecification	An amenity feature (e.g. a characteristic or service) of the Accommodation. This generic

https://schema.org/Place

schema.org/Landform



schema.org		Custom S	Search Q
	Home	Schemas	Documentation

Landform

Thing > Place > Landform

A landform or physical feature. Landform elements include mountains, plains, lakes, rivers, seascape and oceanic waterbody interface features such as bays, peninsulas, seas and so forth, including sub-aqueous terrain features such as submersed mountain ranges, volcanoes, and the great ocean basins.

[more...]

Property	Expected Type	Description
Properties from Place		
additionalProperty	PropertyValue	A property-value pair representing an additional characteristics of the entitity, e.g. a product feature or another characteristic for which there is no matching property in schema.org. Note: Publishers should be aware that applications designed to use specific schema.org properties (e.g. http://schema.org/width, http://schema.org/color, http://schema.org/gtin13,) will typically expect such data to be provided using those properties, rather than using the generic property/value mechanism.
address	PostalAddress or Text	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.

schema.org/BodyOfWater



schema.org	Custom Search
	Home Schemas Documentation
BodyOfWater	

Thing > Place > Landform > BodyOfWater

A body of water, such as a sea, ocean, or lake.

[more...]

Property	Expected Type	Description
Properties from Place		
additionalProperty	PropertyValue	A property-value pair representing an additional characteristics of the entitity, e.g. a product feature or another characteristic for which there is no matching property in schema.org. Note: Publishers should be aware that applications designed to use specific schema.org properties (e.g. http://schema.org/width, http://schema.org/color, http://schema.org/gtin13,) will typically expect such data to be provided using those properties, rather than using the generic property/value mechanism.
address	PostalAddress or Text	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
	LocationFeatureSpecification	An amenity feature (e.g. a characteristic or service) of the Accommodation. This generic

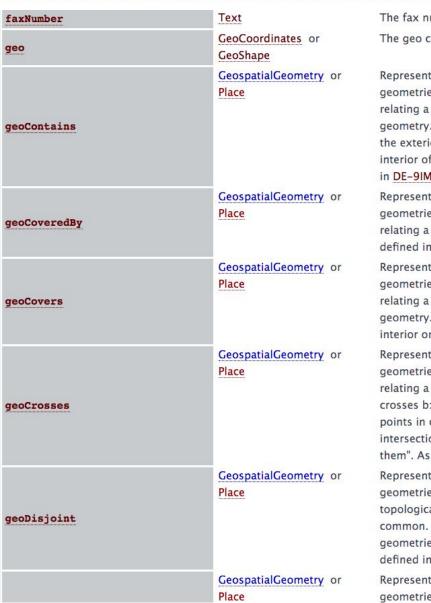
Schema.org Landform Limited Subtypes



More specific Types

- BodyOfWater
- Continent
- Mountain
- Volcano

Schema.org/Place and "Geo" Type



The fax number.

The geo coordinates of the place.

Represents a relationship between two geometries (or the places they represent), relating a containing geometry to a contained geometry. "a contains b iff no points of b lie in the exterior of a, and at least one point of the interior of b lies in the interior of a". As defined in DE-9IM. Represents a relationship between two

geometries (or the places they represent), relating a geometry to another that covers it. As defined in DE-9IM.

Represents a relationship between two geometries (or the places they represent), relating a covering geometry to a covered geometry. "Every point of b is a point of (the interior or boundary of) a". As defined in DE-9IM.

Represents a relationship between two geometries (or the places they represent), relating a geometry to another that crosses it: "a crosses b: they have some but not all interior points in common, and the dimension of the intersection is less than that of at least one of them". As defined in DE-9IM.

Represents spatial relations in which two geometries (or the places they represent) are topologically disjoint: they have no point in common. They form a set of disconnected geometries." (a symmetric relationship, as defined in DE-9IM)

Represents spatial relations in which two geometries (or the places they represent) are



schema.org/geo



geo

Thing > Property > geo

The geo coordinates of the place.

The Geo 'type' relates a 'geometry' value description to a Place

	Lu
Values expected to be one of these types	
GeoCoordinates	
GeoShape	
Used on these types	
Place	
xamples	
xample 1	
Without Markup Microdata RDFa JSON-LD	
xample 2 Without Markup Microdata RDFa JSON-LD	
<script type="application/ld+json"> { "@context": "http://schema.org", "@type": "Place", "geo": {</td><td></td></tr><tr><td>"@type": "GeoCoordinates", "latitude": "40.75",</td><td></td></tr><tr><td>"longitude": "73.98"</td><td></td></tr><tr><td>},</td><td></td></tr><tr><td></td><td></td></tr><tr><td>"name": "Empire State Building"</td><td></td></tr><tr><td>"name": "Empire State Building" }</td><td></td></tr></tbody></table></script>	

schema.org GeoShapes



schema.org	Custom Search
	Home Schemas Documentation

GeoShape

Thing > Intangible > StructuredValue > GeoShape

The geographic shape of a place. A GeoShape can be described using several properties whose values are based on latitude/longitude pairs. Either whitespace or commas can be used to separate latitude and longitude; whitespace should be used when writing a list of several such points.

[more...]

Property	Expected Type	Description
Properties from GeoSha	ipe	
address	PostalAddress or Text	Physical address of the item.
addressCountry	Country or Text	The country. For example, USA. You can also provide the two-letter <u>ISO</u> <u>3166-1 alpha-2 country code</u> .
box	Text	A box is the area enclosed by the rectangle formed by two points. The first point is the lower corner, the second point is the upper corner. A box is expressed as two points separated by a space character.
circle	Text	A circle is the circular region of a specified radius centered at a specified latitude and longitude. A circle is expressed as a pair followed by a radius in meters.
elevation	Number or Text	The elevation of a location (WCS 84). Values may be of the form 'NUMBER UNIT <i>OF</i> MEASUREMENT' (e.g., '1,000 m', '3,200 ft') while numbers alone should be assumed to be a value in meters.
line	Text	A line is a point-to-point path consisting of two or more points. A line is expressed as a series of two or more point objects separated by space.
polygon	Text	A polygon is the area enclosed by a point-to-point path for which the starting and ending points are the same. A polygon is expressed as a series of four or more space delimited points where the first and final



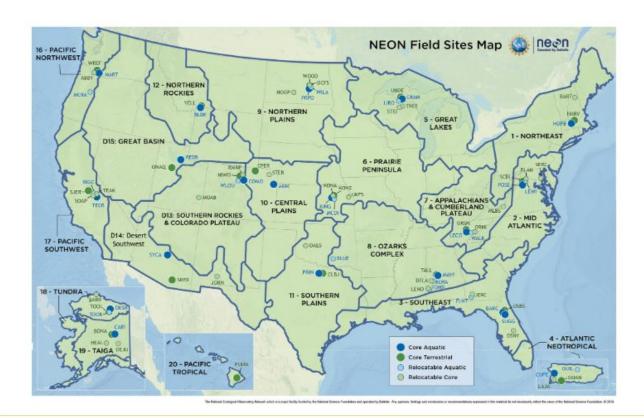
So, why does this matter for "Dataset" search?

Example from NEON



NEON Domains

NEON field sites are placed throughout 20 ecoclimatic domains that represent regions of distinct landforms, vegetation, climate and ecosystem dynamics. The observatory's regional domain approach is designed to statistically represent ecological, physical and biological variability across the continent. Domains range from the Tundra and Taiga in Alaska, to the Atlantic Neotropical in Puerto Rico and the Pacific Neotropical in Hawaii. Data collection methods are also divided between terrestrial field sites and aquatic field sites.



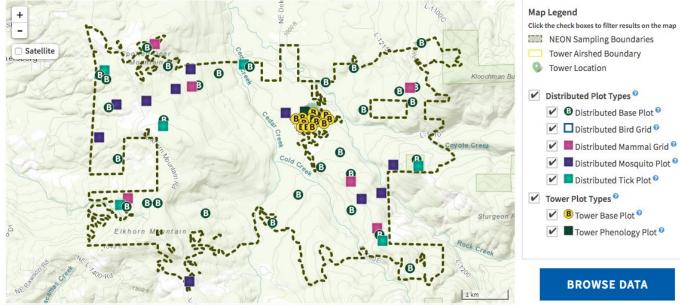
NEON example field-site

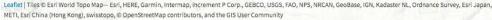


 New Notice Control of the State of the

Abby Road - ABBY

Selocatable Terrestrial | Washington | D16: Pacific Northwest





This map depicts the spatial layout of this field site. Please note that some locations may have moved over time due to logistics, safety and science requirements. This map was updated on April 23, 2019

Rich Information on Landing Pages



Overview

 \sim

The Abby Road (ABBY) field site is located in Yacolt Burn State Forest which is approximately 30 miles from Vancouver, WA and Portland, OR. Yacolt Burn State Forest is a relatively young growth industrial timber production forest. The Washington Department of Natural Resources (DNR) also allows an array of recreational activities including moutain biking, hiking and camping.

Total planned data products for this site: 112

Site Host & Access

~

Site Host:

Washington Department of Natural Resources

Is additional non-NEON research allowed at this site?:

No additional research is allowed in this area at this time. The site hosts have only permitted NEON research activities and are not open to additional research.

Site Characteristics



 \wedge

Site Characteristics

Latitude/Longitude:

45,76243, -122,33033

Elevation: 363 m

Dominant NLCD Classes: Evergreen Forest Grassland/Herbaceous Shrub/Shrub

Mean Annual Temperature: 8C/46.4F

Mean Annual Precipitation: 2530 mm

Sampling plots at ABBY have been established across a mixed landscape of timber plantations ranging from clear cut stands to stands planted in the mid-1960s. Several recreational paths actually cross through the field site.

Site History

Yacolt Burn State Forest is named for the dozens of wildfires that ravaged Clark, Cowlitz and Skamania counties in southwest Washington during September 1902. With no organized system for fighting wildfires, the fires spread across nearly 239,000 acres and caused 38 deaths and widespread property losses. An extended period of hot, dry weather; high wind; an overaccumulation of timber harvest slash; and human carelessness are among the frequently cited causes of these fires. In response, the Washington Legislature established a state fire warden the following year. In 1908, private landowners formed the Washington Fire Protection Association and funded a system of fire wardens and a program of fire prevention on private lands.

The DNR replanted the forests and has maintained the area as working forest in order to:

- · Protect the long-term health of the forest's ecosytems
- Generate revenue for the state
- Provide safe, sustainable, and enjoyable recreational opportunities

Site-specific Topics

- Located in the foothills of the Cascade Mountains in a young forest, ABBY field site provides an
 interesting comparison to NEON's WREF site which is located in a nearby old growth forest that is
 primarily used for research.
- Urban growth west of ABBY since the early 2000s has also brought neighborhoods closer to the forest. This growth has increased the demand for a variety of recreational activities. Urban planning is an active component of the land management of this ecosystem.

Vegetation and Soil

The dominant tree species are Douglas Fir which have been planted by the DNR since the Yakolt Burn in 1902. Many shrubs grow in the forest understory, including salal, Oregon-grape, red and big huckleberries, and west coast rhododendron. Dozens of plant species grow on the forest floor, including queencup beadlily, vanilla leaf, bracken fern, beargrass, twinflower, trillium, and little pipsissewa.

Climate

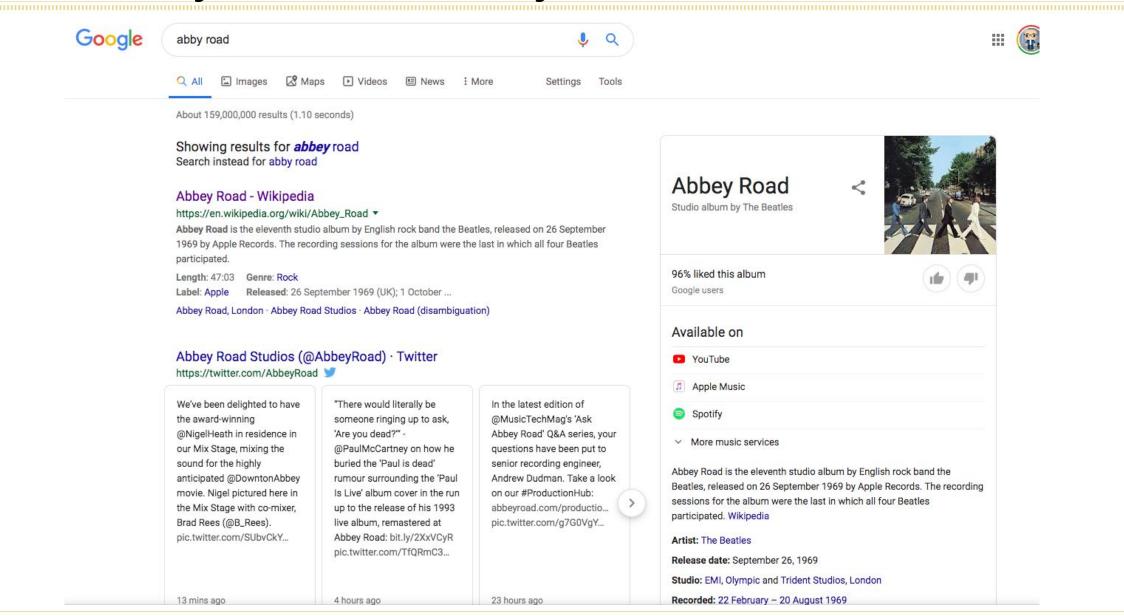
Typical of the Northwest, ABBY experiences very wet falls, winters, and springs; and very dry summers. The nearby Columbia River Gorge affects the valley's climate, contributing to strong

Site characteristics require deeper conceptualization

https://www.neonscience.org/field-sites/field-sites-map/ABBY

Is it Abbey road or Abby Road?







"These aren't the droids you are looking for..."



concepts "These aren't the droids you are looking for..."

Competency Questions?



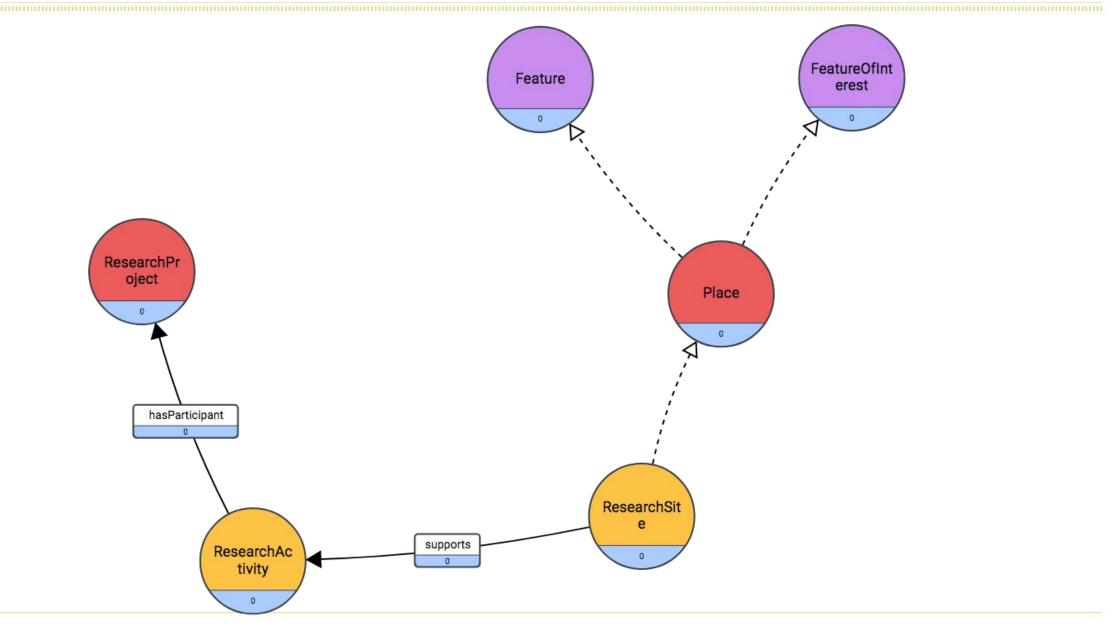
How might a research look for a dataset product?

Questions

- Co-ordinates (various projections and format)
- Country, State/Province
- Ecological region type or Biome
- Land type classification
- Heat zone
- Habitat
- NEON site name
- NEON domain
- NEON 4 letter code
- USGS GNIS feature name (or not GNIS)

Strawman Pattern





schema.org/ResearchProject



Home	Schemas	Documentation	
Home	Schemas	Documentation	

ResearchProject

This term is proposed for full integration into Schema.org, pending implementation feedback and adoption from applications and websites.

Thing > Organization > Project > ResearchProject

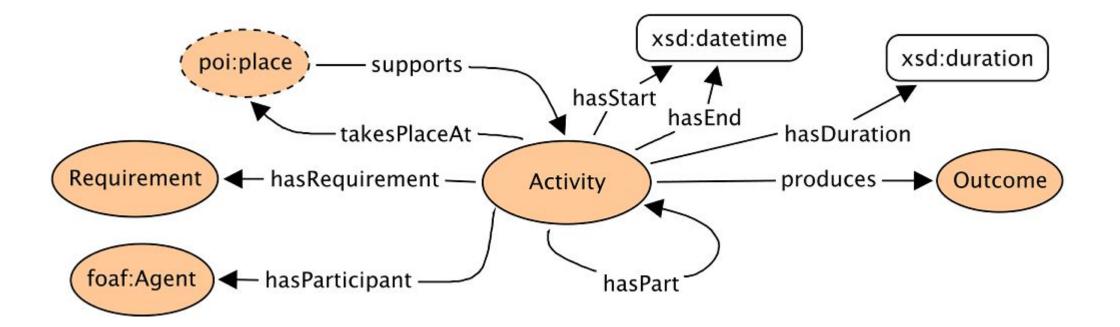
A Research project.

[more...]

Property	Expected Type	Description
Properties from Organization		
actionableFeedbackPolicy	CreativeWork or URL	For a <u>NewsMediaOrganization</u> or other news-related <u>Organization</u> , a statement about public engagement activities (for news media, the newsroom's), including involving the public – digitally or otherwise in coverage decisions, reporting and activities after publication.
address	PostalAddress or Text	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
alumni	Person	Alumni of an organization. Inverse property: <u>alumniOf</u> .
areaServed	AdministrativeArea or <u>GeoShape</u> or <u>Place</u> or <u>Text</u>	The geographic area where a service or offered item is provided. Supersedes <u>serviceArea</u> .
award	Text	An award won by or for this item. Supersedes awards.
brand	Brand or Organization	The brand(s) associated with a product or service, or the brand(s) maintained by an organization or business person.

Activity ODP





http://ontologydesignpatterns.org/wiki/Submissions:An Ontology_Design_Pattern_for_Activity_Reasoning

OGC API - Features



OGC API - Features

This GitHub repository contains OGC's standard for querying geospatial information on the web, "OGC API - Features".

OGC API standards define modular API building blocks to spatially enable Web APIs in a consistent way. OpenAPI is used to define the reusable API building blocks with responses in JSON and HTML.

The OGC API family of standards is organized by resource type. OGC API Features specifies the fundamental API building blocks for interacting with features. The spatial data community uses the term 'feature' for things in the real world that are of interest.

If you are unfamiliar with the term 'feature', the explanations on Spatial Things, Features and Geometry in the W3C/OGC Spatial Data on the Web Best Practice document provide more detail.

Overview

OGC API Features provides access to collections of geospatial data.

GET /collections

Lists the collections of data on the server that can be queried (7.13), and each describes basic information about the geospatial data collection, like its id and description, as well as the spatial and temporal extents of all the data contained

GET /collections/buildings/items?bbox=160.6,-55.95,-170,-25.89

Requests all the data in the collection "buildings" that is in the New Zealand economic zone. The response format (typically HTML or a GeoJSON feature collection, but GML is supported, too, and extensions can easily supply others) is determined using HTTP content negotiation.

Data is returned in pageable chunks, with each response containing a next link as many collections are quite large. The core specification supports a few basic filters, in addition to the bbox filter above, with extensions providing more advanced options (7.15).

GET /collections/{collectionId}/items/{featureId}

Returns a single 'feature' - something in the real-world (a building, a stream, a county, etc.) that typically is described by a geometry plus other properties. This provides a stable, canonical URL to link to the 'thing' (7.16).

https://github.com/opengeospatial/WFS_FES

Spatial Data on the Web



Spatial Data on the Web Best **Practices**

W3C Working Group Note 28 September 2017

This version:

https://www.w3.org/TR/2017/NOTE-sdw-bp-20170928/

Latest published version:

https://www.w3.org/TR/sdw-bp/

Latest editor's draft:

https://w3c.github.io/sdw/bp/

Editors:

Linda van den Brink, Geonovum Payam Barnaghi, University of Surrey

Contributors:

Phil Archer

Jon Blower

Newton Calegari

Byron Cochrane

Simon Cox

François Daoust

Andreas Harth

Bart van Leeuwen

Josh Lieberman

Chris Little

Andy Mabbett

Andrea Perego

Peter Parslow

Ed Parsons



Previous version:

https://www.w3.org/TR/2017/NOTE-sdw-bp-20170511/

Jeremy Tandy, Met Office





-Feature-



NOTE

Although we have borrowed the description of <u>Spatial Thing</u> from [<u>W3C-BASIC-GEO</u>], the formal [<u>RDF-SCHEMA</u>] definition of <u>w3cgeo:SpatialThing</u> doesn't quite suit our purpose as there is the potential for confusion about whether it is *disjoint* from geometry. The definition of <u>geosparql:Feature</u>, which is derived from the [<u>ISO-19109</u>] definition of <u>Feature</u>, is a better semantic fit for <u>Spatial Thing</u> as it is explicitly specified as being disjoint from geosparql:Geometry.

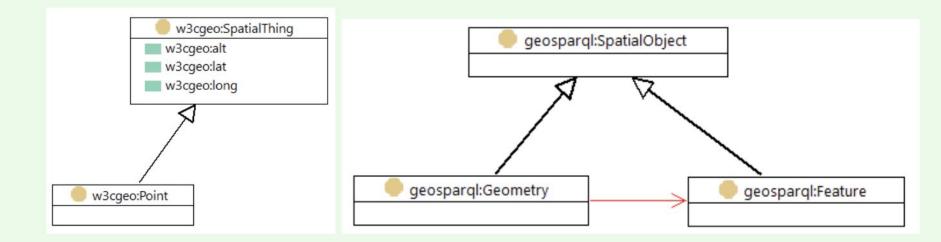


Figure 1 The main classes from [W3C-BASIC-GEO] and [GeoSPARQL] compared - pseudo-UML notation. The red arrow indicates 'disjoint classes'.

Feature of Interest



1. Introduction

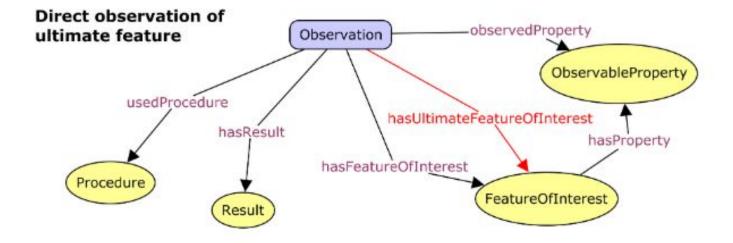
This section is non-normative.

Sensors are a major source of data available on the Web today. While sensor data may be published as mere values, searching, reusing, integrating, and interpreting these data requires more than just the observation results. Of equal importance for the proper interpretation of these values is information about the studied feature of interest, such as a river, the observed property, such as flow velocity, the utilized sampling strategy, such as the specific locations and times at which the velocity was measured, and a variety of other information. OGC's Sensor Web Enablement standards [OandM], [SensorML] provide a means to annotate sensors and their observations. However, these standards are not integrated and aligned with W3C Semantic Web technologies and Linked Data in particular, which are key drivers for creating and maintaining a global and densely interconnected graph of data. With the rise of the Web of Things and smart cities and homes more generally, actuators and the data they produce also become first-class citizens of the Web. Given their close relation to sensors, observations, procedures, and features of interest, it is desirable to provide a common ontology that also includes actuators and actuation. Finally, with the increasing diversity of data and data providers, definitions such as those for sensors need to be broadened, e.g., to include social sensing. The following specifications introduce the new Semantic Sensor Network (SSN) and Sensor, Observation, Sample, and Actuator (SOSA) ontologies that are set out to provide flexible but coherent perspectives for representing the entities, relations, and activities involved in sensing, sampling, and actuation. SOSA provides a lightweight core for SSN and aims at broadening the target audience and application areas that can make use of Semantic Web ontologies. At the same time, SOSA acts as minimal interoperability fall-back level, i.e., it defines those common classes and properties for which data can be safely exchanged across all uses of SSN, its modules, and SOSA.

https://www.w3.org/TR/vocab-ssn/

FOI and Observation





https://www.w3.org/TR/vocab-ssn-ext/

geosparql:Geometry?



NOTE

Although we have borrowed the description of <u>Spatial Thing</u> from [<u>W3C-BASIC-GEO</u>], the formal [<u>RDF-SCHEMA</u>] definition of <u>w3cgeo:SpatialThing</u> doesn't quite suit our purpose as there is the potential for confusion about whether it is *disjoint* from geometry. The definition of <u>geosparql:Feature</u>, which is derived from the [<u>ISO-19109</u>] definition of <u>Feature</u>, is a better semantic fit for <u>Spatial Thing</u> as it is explicitly specified as being disjoint from geosparql:Geometry.

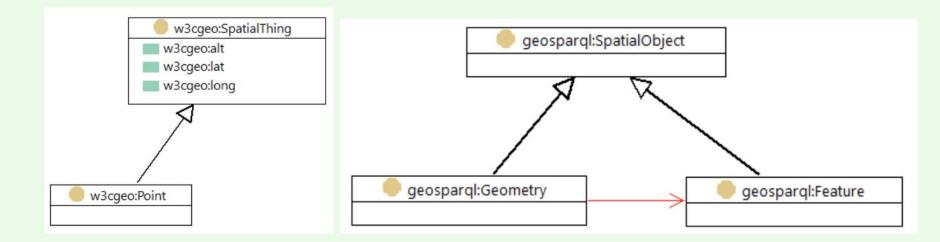


Figure 1 The main classes from [W3C-BASIC-GEO] and [GeoSPARQL] compared - pseudo-UML notation. The red arrow indicates 'disjoint classes'.

GeoJSON(LD), GeoShape or Geosparql?



1

charlesvardeman commented on Jan 28

+ 🙂 🚥

This is a followup from https://github.com/earthcubearchitectureproject418/p418Vocabulary#describing-a-datasets-spatial-coverage topic. Specification of spatial geometries as Linked Data has several approaches, that I think (but could be wrong) are not necessarily compatible. The OGC ELFIE draft technical report states:

"GeoJSON-LD [http://geojson.org/geojson-Id/] was seen as a logical solution, but as described in the outstanding issues for GeoJSON-LD, nested GeoJSON coordinate arrays are not supported by JSON- LD parsers. Additionally, the schema.org geometry schema was seen as under specified in that

among other reasons, it does not provide a default coordinate reference system or a mechanism to declare one. For this reason, the GeoSPARQL well known text format was included providing a precise and JSON-LD compatible geometry format. However, provided the technical limitations listed below, if some logical assumptions are made for the use of schema.org geometry (schema:GeoShape [https://schema.org/GeoShape]), it can be used to satisfy the basic geometry preview use case with significantly lower technical overhead than support for the full well known text standard. Further work should look more closely at this issue in an attempt to reconcile and provide guidance."

I think some of this is resolved in the proposed JSON-LD 1.1 with "bbox": {"@container": "@list"} that creates ordered arrays. The alternative used by GeoSPARQL and several OGC projects is to use WKT string representations which work will with linked data tools (spatial extensions in triple stores) but break web-based toolsets like leaflet and openlayers. Although there are workarounds by adding wkt string to geojson translators in javascript code.

https://github.com/ESIPFed/science-on-schema.org/issues/8

Example from ELFIE



This is only an example and probably not syntatically correct. "@context": ["https://opengeospatial.github.io/ELFIE/json-ld/elf.jsonld", "https://opengeospatial.github.io/ELFIE/json-ld/hyf.jsonld", "gsp": "http://www.opengeospatial.org/standards/geosparql/", "geojson": "https://purl.org/geojson/vocab#", "coordinates": { "@id": "geojson:coordinates", "@container" : "@list", "@values" : { "@type" : "geojson:Coordinate", "@container" : "@set", "@values" : [{"@type" : "xsd:double", "@id":"geo:longitude"}, {"@type" : "xsd:double", "@id":"geo:latitude"} }], "geo": { "@type": "schema:GeoCoordinates", "schema:latitude": 43.2022, "schema:longitude": -89.5302 }, "gsp:hasGeometry": { "@type": "gsp:Geometry", "gsp:asWKT": "POINT (-89.53022 43.20225)" }, "@graph" : [{ "@id" : "geojson:Point", "coordinates" : [-89.53022 43.20225 }]

Alternative Geometry Specification?

```
"@context": "http://schema.org/",
"@type": "Dataset",
"name": "XYZ",
...
"spatialCoverage": {
 "@type": "Place",
 "geo": {
     "@type": "GeoShape",
     ....
   },
   "additionalProperty": [
        "@type": "PropertyValue",
        "additionalType": "http://www.wikidata.org/entity/Q161779",
        "name": "Spatial Reference System for Dataset: XYZ",
        "propertyID": "SRS",
        "value": "http://www.opengis.net/def/crs/OGC/1.3/CRS84",
      },{
        "@type": "PropertyValue",
        "additionalType": "http://www.wikidata.org/entity/Q4018860",
        "name": "Well-Known Text for Dataset: XYZ",
        "propertyID": "WKT",
        "value": "POLYGON ((-75.8183 -68.4817, -68.5033 -68.4817, -68.5033 -65.08, -75.81
```

