HXL Geolocation Standard – *Draft v3*

The Humanitarian eXchange Language geolocation standard (HXL-Geo) specifies how geographic information is to be included into datasets that use the HXL vocabulary. HXL-Geo follows a general principle of HXL in that it is *not* intended to replace any operational systems. Instead, the intention of HXL-Geo is to improve interoperability of existing systems and make the information stored in those systems more accessible.

HXL targets two different levels at which geographic information relevant for humanitarian purposes is collected and processed: Agencies/organizations that have their own (potentially complex) data management systems in place, and small NGOs and implementing partners that work mostly with ad-hoc management tools such as Excel spreadsheets. HXL-Geo needs to complement both user groups by enabling the translation of existing geographic information into HXL, as well as the on-the-fly generation of new geographic information.

The approaches for these two different ways of handling geographic information are described below. They are in line with the Linked Data approach used for HXL. The RDF predicates required for annotating the geographic information are part of the HXL domain vocabulary. [1]

1. Mapping existing geographic information into HXL

Humanitarian organizations and agencies often have a management solution for their geographic information in place. These can range from shared intranet folders through GIS servers and different kinds of registries or catalogues (such as the common operational datasets) to full-fledged spatial data infrastructures. The goal of HXL-Geo is to make these contents more accessible in the context of HXL, as the current situation requires potential users to know beforehand that the respective collection actually exists – otherwise, it is very hard to get hold of the data.

In order to provide these data in a way that is in line with HXL, the minimum requirement is that the metadata for those datasets are made available online according to the Linked Data principles.[2] The first requirement is that a unique URI has to be assigned to the dataset according to the HXL URI pattern guidelines [3]. These guidelines include patterns for the reuse of existing unique identifiers, such as P-Codes, which are unique place and administrative unit identifiers generated by OCHA or national authorities.[3] The HXL domain vocabulary defines the core metadata properties that shall be used to describe the datasets (see the *location* section of the vocabulary) as statements about this URI. Section 3 of this document also lists the corresponding classes and properties.

---

[1] See http://hxl.humanitarianresponse.info for the vocabulary specification.

[2] See https://sites.google.com/site/hxlproject/ for more information about HXL and the underlying Linked Data principles.

[3] See https://sites.google.com/site/ochaimwiki/geodata-preparation-manual/p-code-guidelines for the P-Code guidelines.

One of the statements about such a geodata resource should point to the URL where the original dataset is available for download or accessible through a Web service. HXL-Geo is also capable of handling the fact that this may not be possible for specific datasets (e.g. due to intellectual property right issues). In these cases, the corresponding entry in the HXL registry only consists of the metadata, without a direct pointer to the data. In this case, the HXL metadata entry created for the dataset *must* contain some information about how to obtain the data.

The most relevant example of this metadata-only approach is the current implementation of the Common Operational Datasets (CODs). [4] The CODs for administrative boundaries and populated places serve as the common standard of reference for locations in humanitarian responses and provide unique identifiers (p-codes) for each administrative boundary and populated place. Currently, these datasets are typically provided as downloadable GIS data, but the geometry and attributes of a given feature are not directly web-accessible. However, HXL statements can still refer to these features through the URL to the downloadable resource.

2. On-the-fly generation of new geographic information

For any organizations that cannot or do not want to host geographic information themselves, HXL-Geo offers the RDF properties required to embed the geography of the reported features. For these cases, HXL-Geo makes a trade-off between complexity of the data model and expressivity (i.e., what kinds of geometries can be stored). Generators of on-the-fly geo-information for HXL are unlikely to create highly complex geographic information. We therefore opted for the Well-Known Text format (WKT), a straightforward solution that is both standardized [1,9] and very compact in storage. The WKT standard also includes a pointer to the spatial reference system (SRS) that the geodata use, so that no separate handling of the SRS is necessary. Moreover, WKT is easy to export into other standard encodings for geographic information, such as GML [4], KML [5], or ESRI file formats (Shape or GeoDatabase).[5]

WKT also prepares HXL to offer spatial queries directly from the HXL registry. The registry can be queried using the SPARQL query language; using WKT facilitates an extension of the registry with GeoSPARQL in the future [2].

Since reporting organizations cannot be expected to write HXL-Geo code "by hand", the ReportWhere tool [6] will contain a mapping functionality that allows them to easily create geometries and metadata for their datasets. The WKT encoding can also be used to provide a generic idea of the geometries represented in existing geodatasets hosted externally. A typical use case includes the specification of a bounding box for the referenced dataset. Providing this information enables simple spatial queries (e.g. for all geographic information available within a specific area) and is therefore strongly

[4] http://cod.humanitarianresponse.info

[5] Organizations that want to import HXL data into their systems with geographic information encoded in any of these standards need to implement the translation from WKT into the desired format in their HXL adapter. This export functionality is not part of HXL.

encouraged. HXL adapters that export existing geodata should implement this functionality.

3. Recommended classes and predicates for data annotation

The modeling approach taken in HXL is based on RDF/Linked Data, where statements about things can be made in the form *Subject – Predicate – Object* (e.g. *OCHA – is a – Organization*; see [8] for an introduction on Linked Data). The classes listed below define the types of things that can be used as subjects and objects in such triples in the context of HXL-Geo. The predicates define the different relationships that can be used to link them or define attribute values for them.

The classes and predicates defined in the following are based on the OGC Simple Feature Model and reduced to a subset of the full model for simplicity and ease of use (see Figure 1). These lists only form the core vocabulary of HXL-Geo and can be extended by organizations if need be. They also strongly reflect the classes and predicates proposed in the OGC GeoSPARQL draft and may be replaced by those in a future version of HXL. We opted to replicate them within HXL because (a) GeoSPARQL is not standardized yet, and (b) the classes and predicates defined in the draft GeoSPARQL spec are not yet identified by resolvable URIs.

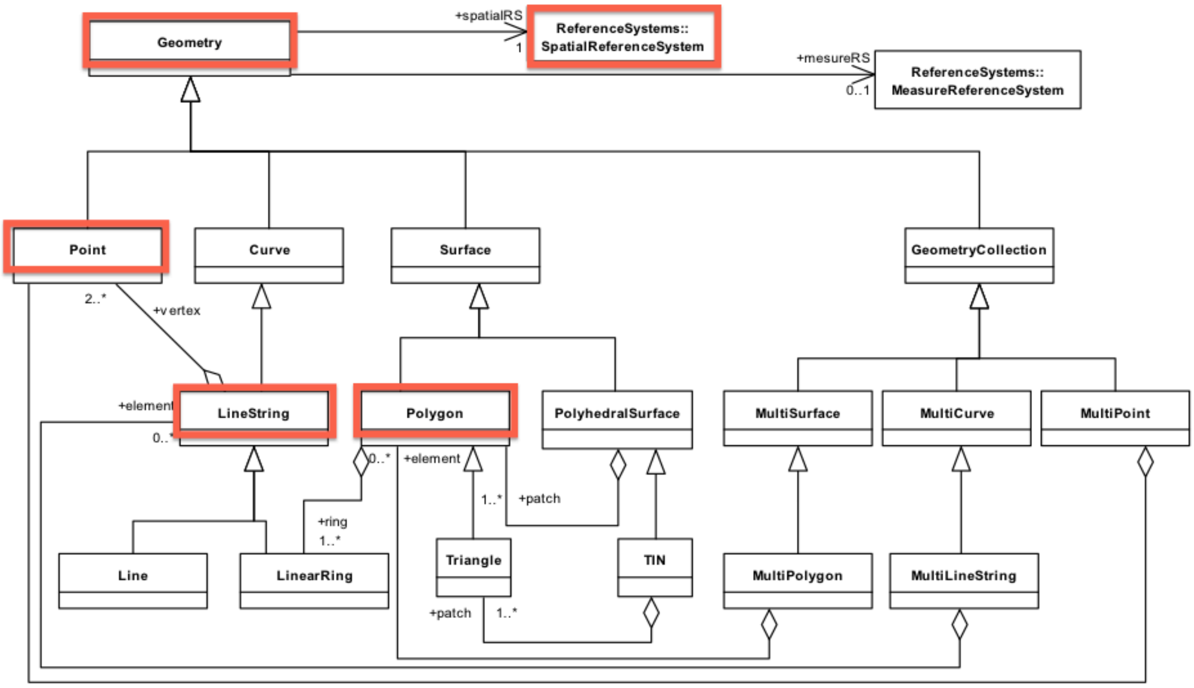


Figure 1: OGC Simple Feature Geometry Class Hierarchy [9]. The classes highlighted in red are reflected in HXL-Geo.

1. Classes

The classes specify which types of things can occur as subjects and objects in a HXL statement.

hxl:Feature

Description	Any real-word phenomenon with spatial extent, i.e. size, shape, or position. The vocabulary document at http://hxl.humnitarianresponse.info will be incrementally extended with specific feature types such as administrative units.
URI	http://hxl.humanitarianresponse.info#Feature
Superclass	http://www.w3.org/2003/01/geo/wgs84_pos#SpatialThing

hxl:Geometry

Description	Abstract root class of the geometry. This class should not be instantiated; use any of the subclasses when creating actual geometries.
URI	http://hxl.humanitarianresponse.info#Geometry

hxl:Point

Description	A Point is a 0-dimensional geometric object and represents a single location in a two- or three-dimensional coordinate space.
URI	http://hxl.humanitarianresponse.info - Point
Superclass	hxl:Geometry

hxl:LineString

Description	A LineString is a Curve with linear interpolation between Points. Each consecutive pair of Points defines a Line segment.
URI	http://hxl.humanitarianresponse.info#LineString
Superclass	hxl:Geometry

hxl:Polygon

Description	A Polygon is a planar Surface defined by 1 exterior boundary and 0 or more interior boundaries. Each interior boundary defines a hole in the Polygon.
URI	http://hxl.humanitarianresponse.info#Polygon

Superclass	hxl:Geometry

hxl:BoundingBox

Description	A BoundingBox is a Polygon defined by 1 exterior boundary and 0 interior boundaries. The exterior boundary consists of 4 points arranged as a rectangle.
URI	http://hxl.humanitarianresponse.info#BoundingBox
Superclass	hxl:Polygon

2. Predicates

The predicate definitions specify between which kinds of things these predicates can be used, i.e., what kinds of subjects and objects can be part of a statement that uses this predicate. The *domain* defines the types of subjects; the *range* defines the types of objects. The topological relation predicates are strongly influenced by the Simple Feature Query Functions [9], but contain additional, less rigorously defined predicates (e.g. *hxl:near*) to meet the requirements of the domain. Being able to make topological statements about the geometry of a feature is especially important in the context of HXL to be able to roughly locate features with an undefined geometry – e.g., an organization may want to report mines *within* a certain area, without being forced to specify a geometry for the minefield if its exact extent is unknown. For generic metadata statements – such as descriptions, point of contact, etc. – HXL uses the Dublin Core Metadata Terms [10]. The most important ones are listed here for completeness with a note on intended use within HXL.

hxl:hasGeometry

Description	Relates a feature to its geometry.
URI	http://hxl.humanitarianresponse.info#hasGeometry
Domain	hxl:Feature
Range	hxl:Geometry

hxl:asWKT

Description	A geometry serialized as WKT. Note that the WKT serialization contains the spatial reference system (SRS), as specified in [9]. If no SRS is given, HXL uses EPSG 4326 as the standard SRS. This is the standard SRS used by GPS, based on the WGS 84 datum.
URI	http://hxl.humanitarianresponse.info#asWKT

Domain	hxl:Geometry
Range	rdfs:Literal

hxl:equals

Description	Topological relation. States that two geometries are equal.
URI	http://hxl.humanitarianresponse.info#equals
Domain	hxl:Geometry
Range	hxl:Geometry

hxl:intersects

Description	Topological relation. States that one geometry intersects another one.
URI	http://hxl.humanitarianresponse.info#intersects
Domain	hxl:Geometry
Range	hxl:Geometry

hxl:touches

Description	Topological relation. States that one geometry touches another one.
URI	http://hxl.humanitarianresponse.info#touches
Domain	hxl:Geometry
Range	hxl:Geometry

hxl:crosses

Description	Topological relation. States that one geometry crosses another one.
URI	http://hxl.humanitarianresponse.info#crosses
Domain	hxl:Geometry
Range	hxl:Geometry

hxl:within

Description	Topological relation. States that one geometry is within another one.

URI	http://hxl.humanitarianresponse.info#within
Domain	hxl:Geometry
Range	hxl:Geometry
Inverse	hxl:contains

hxl:contains

Description	Topological relation. States that one geometry contains another one.
URI	http://hxl.humanitarianresponse.info#contains
Domain	hxl:Geometry
Range	hxl:Geometry
Inverse	Hxl:within

hxl:overlaps

Description	Topological relation. States that two geometries overlap.
URI	http://hxl.humanitarianresponse.info#overlaps
Domain	hxl:Geometry
Range	hxl:Geometry

hxl:near

Description	Topological relation. States that one geometry is near another one.
URI	http://hxl.humanitarianresponse.info#near
Domain	hxl:Geometry
Range	hxl:Geometry

Dublin Core Metadata Terms:

dc:rightsHolder

Description	„*A person or organization owning or managing rights over the resource*“ [10]. Indicates the point of contact for datasets that are not available online.

URI	http://purl.org/dc/terms/rightsHolder

dc:spatial

Description	*„Spatial characteristics of the resource“* [10]. Can be used to refer to the bounding box of an externally stored dataset.
URI	http://purl.org/dc/terms/spatial

dc:subject

Description	*„Typically, the subject will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary. To describe the spatial or temporal topic of the resource, use the Coverage element.“* [10].
URI	http://purl.org/dc/terms/subject

dc:format

Description	Indicate the file format or service type the user can expect for externally stored datasets.
URI	http://purl.org/dc/terms/format

dc:source

Description	The URL at which an externally stored dataset is available.
URI	http://purl.org/dc/terms/source

4. Examples

The following examples in Turtle notation[6] give an overview of how to refer to an existing dataset using HXL-Geo (1), and how encode geometry directly in a HXL file (2). Note that the given URIs are *not* referring to any actual content and only assigned for illustration. The HXL URI Pattern Guidelines [5] will specify how these URIs have to be constructed. The examples below use the following namespace prefixes:

[6] See http://www.w3.org/TeamSubmission/turtle/.

```
@prefix hxl:  <http://carsten.io/hxl/ns#> .
@prefix rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dc:   <http://purl.org/dc/terms/> .
@prefix xsd:  <http://www.w3.org/2001/XMLSchema#> .
```

1. Existing dataset hosted in a separate organization-specific system

This example encodes a P-Code dataset in HXL. It links to the COD dataset that contains the geometry for the Kenyan province *Eastern* (pcode = KE4) and provides the bounding box for this province as a rough spatial reference.

```
<http://exam.pl/resource/geolocation/pcode/KE4>
   rdf:type    hxl:Feature ;
   dc:title    "Eastern"^^xsd:string ;
   dc:spatial <http://exam.pl/resource/geolocation/pcode/KE4/bbox> ;
   dc:source   <http://cod.humanitarianresponse.info/sites/default/
files/ken_provinces.zip> ;
   dc:format   "ESRI Shape file, zipped."^^xsd:string .

<http://exam.pl/resource/geolocation/pcode/KE4/bbox>
   rdf:type    hxl:BoundingBox ;
   hxl:asWKT   """

        <http://www.opengis.net/def/crs/OGC/1.3/CRS84>
        Polygon((36.1 4.5, 39.4 4.5, 39.4 -3.1, 36.1 -3.1, 36.1 4.5))

   """^^<http://www.opengis.net/def/dataType/OGC-SF/1.0/WKTLiteral> .
```

2. New dataset with complete storage in HXL, including geometry

This example encodes the location of the Nairobi Airport in WKT.

```
<http://exam.pl/resource/geolocation/airport/1234>
      rdf:type    hxl:Feature ;
      dc:title    "Nairobi Airport"^^xsd:string ;
      hxl:hasGeometry
<http://exam.pl/resource/geolocation/airport/1234/geometry>  .

<http://exam.pl/resource/geolocation/airport/1234/geometry>
      rdf:type hxl:Point ;
      hxl:asWKT """

        <http://www.opengis.net/def/crs/OGC/1.3/CRS84>
```

```
        Point((36.947510 -1.312540))

   """^^<http://www.opengis.net/def/dataType/OGC-SF/1.0/WKTLiteral> .
```

3. “Fuzzy” spatial referencing using topological relations

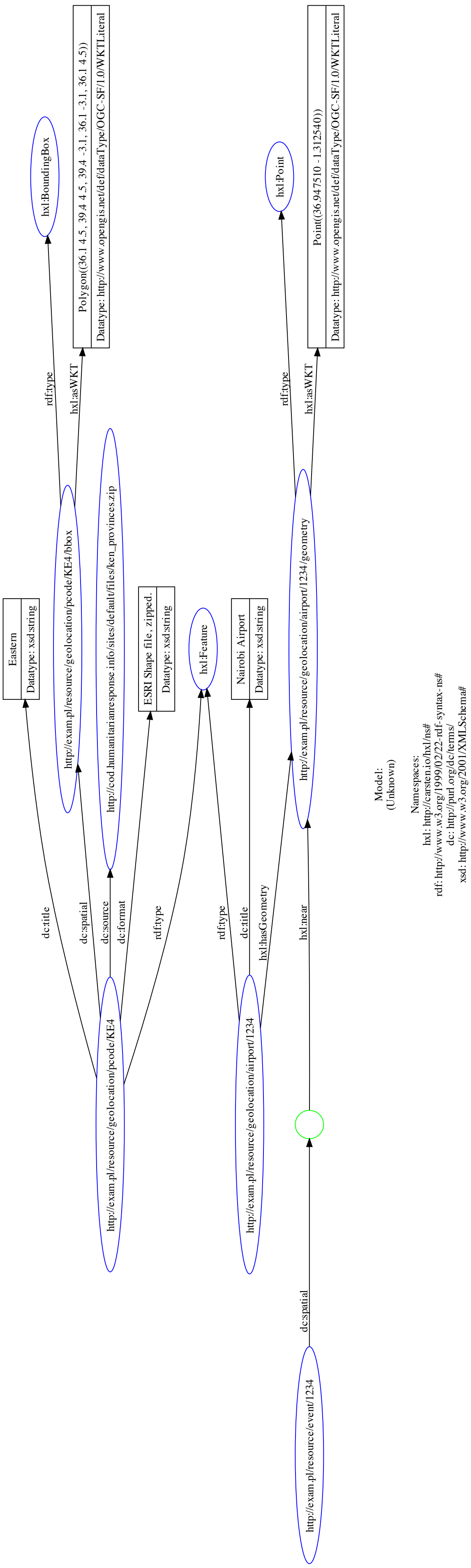
This example reports an event where the exact location of the event is unknown, and therefore represented as a blank node (_:b0); it is only known that the event happened near Nairobi Airport.

```
<http://exam.pl/resource/event/1234>
      dc:spatial _:b0 .

_:b0 hxl:near <http://exam.pl/resource/geolocation/airport/1234/geometry> .
```

4. Graph visualization

Figure 2 gives a visual overview of the examples introduced above.


hxl:BoundingBox
Polygon((36.1 4.5, 39.4 4.5, 39.4 -3.1, 36.1 -3.1, 36.1 4.5))
Datatype: http://www.opengis.net/def/dataType/OGC-SF/1.0/WKTLiteral
hxl:Point
Point(36.947510 -1.312540))
Datatype: http://www.opengis.net/def/dataType/OGC-SF/1.0/WKTLiteral
rdf:type
hxl:asWKT
rdf:type
hxl:asWKT
Eastern
Datatype: xsd:string
http://exam.pl/resource/geolocation/pcode/KE4/bbox
http://cod.humanitarianresponse.info/sites/default/files/ken_provinces.zip
ESRI Shape file, zipped.
Datatype: xsd:string
hxl:Feature
Nairobi Airport
Datatype: xsd:string
http://exam.pl/resource/geolocation/airport/1234/geometry
Model:
(Unknown)
Namespaces:
hxl: http://carsten.io/hxl/ns#
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
dc: http://purl.org/dc/terms/
xsd: http://www.w3.org/2001/XMLSchema#
dc:title
dc:spatial
dc:source
dc:format
rdf:type
rdf:type
dc:title
hxl:hasGeometry
hxl:near
http://exam.pl/resource/geolocation/pcode/KE4
http://exam.pl/resource/geolocation/airport/1234
dc:spatial
http://exam.pl/resource/event/1234

5. References

[1] International Standards Organization (2005) ISO 19125-1:2004. Geographic information – Simple feature access – Part 1: Common architecture. Available from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40114

[2] Open Geospatial Consortium (2011) OGC GeoSPARQL – A Geographic Query Language for RDF Data (11-052r3). Request for comments, available from http://www.opengeospatial.org/standards/requests/80

[3] HXL URI pattern guidelines – tbd.

[4] Open Geospatial Consortium (2007) OpenGIS Geography Markup Language (GML) Encoding Standard. OGC standard available from http://portal.opengeospatial.org/files/?artifact_id=20509

[5] Open Geospatial Consortium (2008) OGC KML. OGC standard available from http://portal.opengeospatial.org/files/?artifact_id=27810

[6] ReportWhere Spec – tbd.

[7] W3C (2012) SPARQL 1.1 Query LanguageW3C. Working Draft, 05 January 2012, available from http://www.w3.org/TR/2012/WD-sparql11-query-20120105/

[8] Tom Heath and Christian Bizer (2011) Linked Data – Evolving the Web into a Global Data Space

[9] OGC (2011) OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture. Version 1.2.1. OGC standard available from http://portal.opengeospatial.org/files/?artifact_id=25355

[10] Dublin Core Metadata Initiative (2010) DCMI Metadata Terms. Vocabulary specification available from http://dublincore.org/documents/dcmi-terms/