[DM-07-058] Geospatial Content Standards

Abstract

Geospatial content Standards for digital data build upon centuries of work in map making and cartography. Since 1994, two organizations have been the primary drivers of these Standards: the Open Geospatial Consortium and ISO / TC 211 Geographic Information / Geomatics. This topic first defines the landscape of Standards development and its objectives. The authority of the Standards bodies is also explained. Finally, geospatial content Standards are divided into five categories, with examples provided for each: (1) feature data; (2) 3D geometry and streaming; (3) rasters and coverages; (4) datacubes, and (5) metadata.

Keywords: data standards, geospatial interoperability, interoperability, metadata, Open Geospatial Consortium

Author & citation

Simmons, S. (2025). Geospatial Content Standards. The Geographic Information Science and Technology Body of Knowledge (Issue 2, 2025 Edition), John P. Wilson (Ed.).
DOI: 10.22224/gistbok/2025.2.10.

Explanation

1. Geospatial Standards Landscape
2. Purpose of Standards
3. Authority of Standards
4. Geospatial Content Standards

1. Geospatial Standards Landscape

Standards define a convention on how to make or do something. Consider a paper map, the symbols and colors for a specific series of maps (such as the topographic quadrangles issued by the USGS) are standardized across the whole series. Concepts such as using the color blue to represent water data back to medieval times or even earlier. Most of these standards are defined by specific publishers and not necessarily shared as paper maps do not need to interoperate.

The use of digital spatial data management, electronic mapping, web presentation of map data, etc. necessitated a change in behavior. The nature of electronic data means that sharing or conflating data without great pain requires common formats and well-defined structures of those data. Further, posting map data to a web browser requires that the data projection be defined, the data structure support use in a web interface, and the browser understand what it is trying to show.

The increasing use of information technology to manage and disseminate geospatial data led to the formation in 1994 of both the Open Geospatial Consortium (OGC: www.ogc.org) and ISO / TC 211: Geographic Information / Geomatics

(https://committee.iso.org/home/tc211). Both organizations were established to define common conventions for managing digital geospatial data.

While both organizations were formed in the same year and continue to closely cooperate to this very day, there are some differences in their missions. ISO focuses on foundational Standards, primarily of a conceptual nature, to define how basic geospatial encodings and operations should be structured. OGC focuses more on implementation Standards – specific encodings of geospatial data types in JSON, XML, etc. and service interfaces to get data to the user.

2. Purpose of Standards

Standards form the foundation upon which to build capabilities, be those data infrastructures, software tools, or downstream applications. Standards define what must be consistent and if ideally designed, only the required aspects of consistency.

An analogy for what Standards once did and now do comes from the world of industrial machinery. The earliest Standards were developed as part of the Industrial Revolution to provide definition for reusable parts, such as bolts. These Standards defined parameters such as diameter, thread pitch, and head dimensions. Standardized bolts and other parts are the fundamental elements used to create machinery. The machinery actually provides functionality that someone can use.

Geospatial fundamental Standards describe how to encode imagery, features, volumetric information, and more. There are also Standards that help describe the data (e.g., ISO 19115 metadata). Collectively, these Standards define the parts that are used to build datasets and interoperability that can be implemented, providing functionality that someone can use.

Machinery needs additional Standardized supporting components to work. Lubricants ensure that the machinery operates smoothly and remains functional for many years. Modern OGC Standards are much like the grease in a machine – they enable other Standardized parts or services to work together and more smoothly. In short... the community knows how to make the machines, now they want to make the machines run better.

Once Standards are established, the user community can implement solutions directly against those Standards or adapt the Standards via profiling (subsetting), specialization, and/or extension to create specifications for their specific use cases.

3. Authority of Standards

OGC and ISO produce international consensus-based Standards and supporting deliverables based on diverse use cases tested in real-world experimentation.

International: requires membership from across the world. Diversity in stakeholders is also key: representation from all parts of the world are necessary for Standards

Development Organizations (SDOs) to be truly reflective of the community and its various levels of maturity.

Consensus (by membership makeup): representative participation by all community segments at a variety of sizes of operation – commercial, government, research/academia. This is a key differentiator for OGC: very few industry associations or standards bodies are so balanced in participation – thus, the work is more likely to be relevant.

Consensus (by process): OGC follows a very open review and approval process:

- open access by members to all parts of the output development process;
- public review of Standards and related work before approval;
- open voting by all members in Standards Working Groups (SWGs)
- open voting by Voting members in final approval; and
- approval by visible vote requiring super-majority agreement.

International and consensus (by process) are required to be considered a de jure Standard in most nations and for contributions to ISO and its national bodies.

4. Geospatial Content Standards

Note that this section does not provide an exhaustive list of geospatial content Standards, but does focus on those most used or fundamental. Geospatial Content Standards exist in a large architecture of Standards for collection, encoding, and dissemination of geospatial data, such as illustrated in Figure 1.

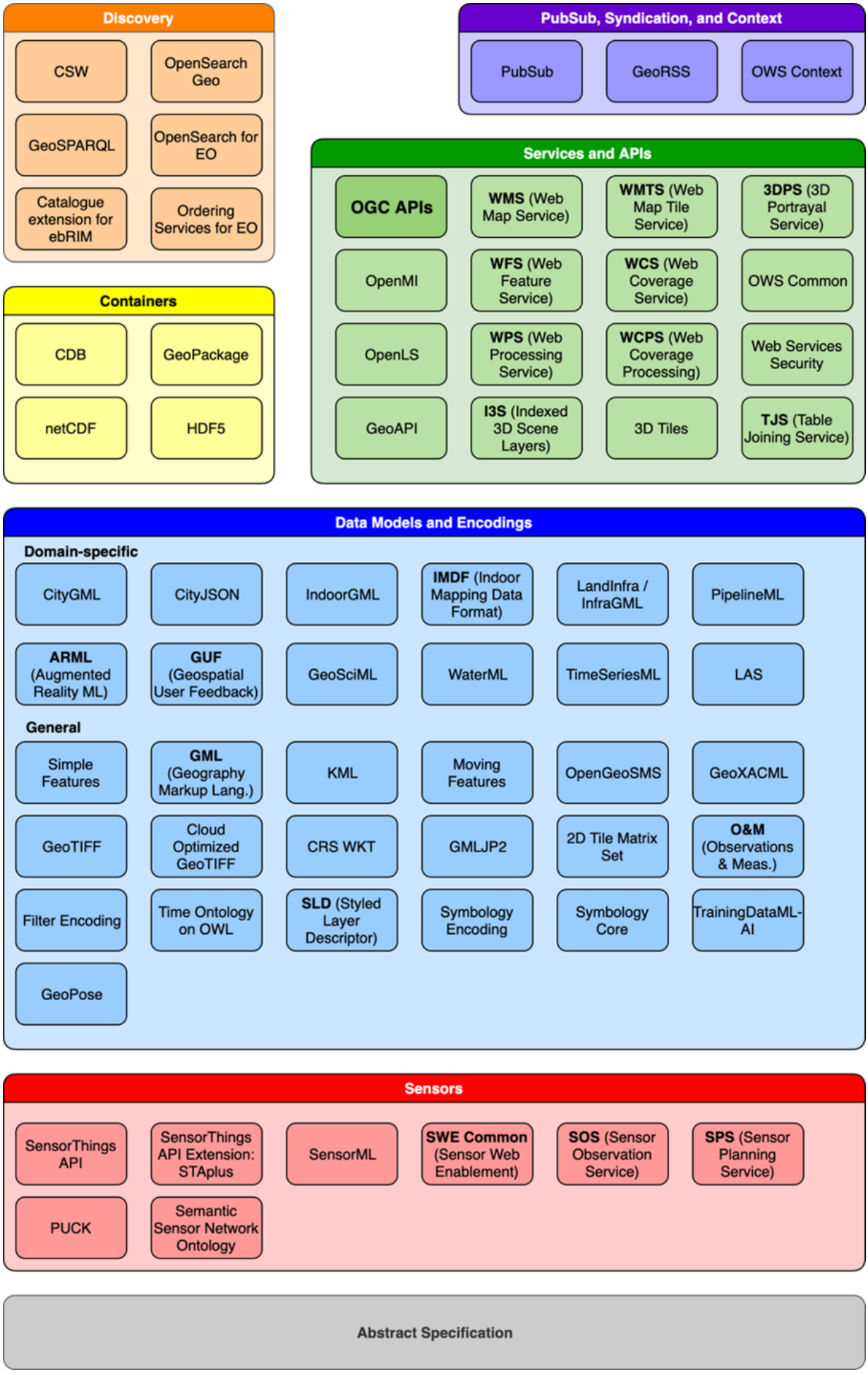

Discovery
CSW
OpenSearch Geo
GeoSPARQL
OpenSearch for EO
Catalogue extension for ebRIM
Ordering Services for EO
PubSub, Syndication, and Context
PubSub
GeoRSS
OWS Context
Containers
CDB
GeoPackage
netCDF
HDF5
Services and APIs
OGC APIs
WMS (Web Map Service)
WMTS (Web Map Tile Service)
3DPS (3D Portrayal Service)
OpenMI
WFS (Web Feature Service)
WCS (Web Coverage Service)
OWS Common
OpenLS
WPS (Web Processing Service)
WCPS (Web Coverage Processing)
Web Services Security
GeoAPI
I3S (Indexed 3D Scene Layers)
3D Tiles
TJS (Table Joining Service)
Data Models and Encodings
Domain-specific
CityGML
CityJSON
IndoorGML
IMDF (Indoor Mapping Data Format)
LandInfra / InfraGML
PipelineML
ARML (Augmented Reality ML)
GUF (Geospatial User Feedback)
GeoSciML
WaterML
TimeSeriesML
LAS
General
Simple Features
GML (Geography Markup Lang.)
KML
Moving Features
OpenGeoSMS
GeoXACML
GeoTIFF
Cloud Optimized GeoTIFF
CRS WKT
GMLJP2
2D Tile Matrix Set
O&M (Observations & Meas.)
Filter Encoding
Time Ontology on OWL
SLD (Styled Layer Descriptor)
Symbology Encoding
Symbology Core
TrainingDataML-AI
GeoPose
Sensors
SensorThings API
SensorThings API Extension: STAplus
SensorML
SWE Common (Sensor Web Enablement)
SOS (Sensor Observation Service)
SPS (Sensor Planning Service)
PUCK
Semantic Sensor Network Ontology
Abstract Specification

Figure 1. The OGC Standard Architecture (2024). Source: OGC.

How to categorize “content” is itself somewhat controversial as there is a continuum across the representation of many types of geospatial data. For the purposes of this entry, four categories will be used: (1) feature data; (2) 3D geometry and streaming; (3) rasters and coverages; and (4) datacubes.

4.1 Feature Data

Feature data is a term synonymous with vector data. This is a representation of points, lines, polygons, curves, and surfaces. Feature data can be 2D or 3D, although there are specialized Standards for 3D content that will be discussed later in this section. Below are some examples of feature data Standards.

Simple Features: OpenGIS® (the original name of OGC) Simple Features Access (SFA) [OGC 06-103r4], also called ISO 19125, describes the common architecture for simple feature geometry. The simple feature geometry object model is computing platform-neutral and is documented in UML. The base Geometry class has subclasses for Point, Curve, Surface and GeometryCollection. Each geometric object is associated with a Spatial Reference System, which describes the coordinate space in which the geometric object is defined. Simple Features is ubiquitous, found in every relational database and it underlies almost overy other feature geometry encoding in use today.

Geography Markup Language (GML): The OpenGIS® Geography Markup Language Encoding Standard [OGC 07-036r1], also ISO 19136, is an XML grammar for expressing geographical features. GML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. As with most XML based grammars, there are two parts to the grammar – the schema that describes the document and the instance document that contains the actual data. A GML document is described using a GML Schema. This allows users and developers to describe generic geographic data sets that contain points, lines and polygons. However, the developers of GML envision communities working to define community-specific application schemas that are specialized extensions of GML. Using application schemas, users can refer to roads, highways, and bridges instead of points, lines and polygons. If everyone in a community agrees to use the same schemas they can exchange data easily and be sure that a road is still a road when they view it.

GeoJSON: Published by the Internet Engineering Task Force (IETF) [RFC 7946] and per that document, “GeoJSON is a geospatial data interchange format based on JavaScript Object Notation (JSON). It defines several types of JSON objects and the manner in which they are combined to represent data about geographic features, their properties, and their spatial extents. GeoJSON uses a geographic coordinate reference system, World Geodetic System 1984, and units of decimal degrees.”

Geoscience Markup Language (GeoSciML): An example of a domain-specific feature data Standard, GeoSciML [OGC 16-008] is a model of geological features commonly described and portrayed in geological maps, cross sections, geological reports and databases. This Standard describes a logical model and GML/XML encoding rules for the exchange of geological map data, geological time scales, boreholes, and metadata for laboratory analyses.

4.2 3D Geometry and Streaming

While most 3D geometry Standards in the community represent the information with feature models, the visualization and analysis of 3D information is more complex than that of 2D. Further 3D voxels may be used to represent volumetric data.

City Geography Markup Language (CityGML): CityGML [OGC 20-010] is an open data model and XML-based format for the storage and exchange of virtual 3D city models. It is an application schema for GML. The aim of the development of CityGML is to reach a common definition of the basic entities, attributes, and relations of a 3D city model. This is especially important with respect to the cost-effective sustainable maintenance of 3D city models, allowing the reuse of the same data in different application fields.

3D Tiles: 3D Tiles [OGC 22-025r4] is an OGC Community Standard developed and maintained by Bentley Systems (formerly Cesium). 3D Tiles is designed for streaming and rendering massive 3D geospatial content such as Photogrammetry, 3D Buildings, BIM/CAD, Instanced Features, and Point Clouds. It defines a hierarchical data structure and a set of tile formats which deliver renderable content. 3D Tiles does not define explicit rules for visualization of the content; a client may visualize 3D Tiles data however it sees fit. A 3D Tiles data set, called a tileset, contains any combination of tile formats organized into a spatial data structure. 3D Tiles are declarative, extendable, and applicable to various types of 3D data. This document also describes 3D Tile Styles, a declarative styling specification which may be applied to tilesets.

Indexed 3D Scene Layers (I3S): I3S [OGC 17-014r9] is a container for arbitrarily large amounts of heterogeneously distributed 3D geographic data. Scene Layers are designed to be used in mobile, desktop, and server-based workflows and can be accessed over the web or as local files. The delivery format and persistence model for Scene Layers, referred to as Indexed 3d Scene Layer (I3S) and Scene Layer Package (SLPK) respectively, are specified in detail in this OGC Community Standard. Both formats are encoded using JSON and binary ArrayBuffers (ECMAScript 2015). I3S is designed to be cloud, web and mobile friendly. I3S is based on JSON, REST and modern web standards and is easy to handle, efficiently parse and render by Web and Mobile Clients. I3S is designed to stream large 3d datasets and is designed for performance and scalability. I3S is designed to support 3D geospatial content and supports the requisite coordinate reference systems and height models in conjunction with a rich set of layer types.

4.3 Rasters and Coverages

Rasters (gridded data such as satellite imagery) are a subset of a data type known as coverages. Coverages are digital representations of space-time varying phenomena. In practice, coverages encompass regular and irregular grids, point clouds, and general meshes.

GeoTIFF: The OGC GeoTIFF Standard [OGC 19-008r4] defines the Geographic Tagged Image File Format (GeoTIFF) by specifying requirements and encoding rules for using the Tagged Image File Format (TIFF) for the exchange of georeferenced or geocoded imagery. The GeoTIFF 1.1 standard formalizes the existing community GeoTIFF specification version 1.0 and aligns it with the continuing addition of data to the EPSG Geodetic Parameter Dataset. The GeoTIFF format is used throughout the geospatial and earth science communities to share geographic image data.

Cloud Optimized GeoTIFF (COG): COG [OGC 21-026] relies on two characteristics of the TIFF v6 format (tiles and reduced resolution subfiles), GeoTIFF keys for georeference, and the HTTP range, which allows for efficient downloading of parts of imagery and grid coverage data on the web and to make fast data visualization of TIFF or BigTIFF files and fast geospatial processing workflows possible. COG-aware applications can download only the information they need to visualize or process the data on the web. Numerous remote sensing datasets are available in cloud storage facilities that can benefit from optimized visualization and processing. This standard formalizes the requirements for a TIFF file to become a COG file and for the HTTP server to make COG files available in a fast fashion on the web.

Coverage Implementation Schema (CIS): CIS [OGC 09-146r8] specifies the OGC coverage model by establishing a concrete, interoperable, conformance-testable coverage structure. It is based on the abstract concepts of OGC Abstract Topic 6 (which is identical to ISO 19123) which specifies an abstract model which is not per se interoperable – in other words, many different and incompatible implementations of the abstract model are possible. CIS, on the other hand, is interoperable in the sense that coverages can be conformance tested, regardless of their data format encoding, down to the level of single “pixels” or “voxels.” Coverages can be encoded in any suitable format (such as GML, JSON, GeoTIFF, or NetCDF) and can be partitioned, e.g., for a time-interleaved representation.

4.4 Data Cubes

Datacubes are multidimensional collections of data, often gridded and having at least two spatial dimensions. Datacubes are conceptually generally encoded as coverages. However, multidimensional data often have characteristics that require specific data stores to improve performance of retrieval and analysis of data and these content models may not perfectly adhere to a coverage concept.

Hierarchical Data Format (HDF) 5: HDF5 [OGC 18-043r3] is a data model, a programming interface, and a storage model for keeping and managing data. It supports an unlimited variety of data types, and is designed to be flexible and efficient for large and complex data. HDF5 is extensible via customizing data types, allowing communities and their applications to evolve in the use of HDF5. HDF5 can be used to encode netCDF (below).

Network Common Data Form (netCDF): NetCDF [OGC 10-090r3] is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data. Conventions define metadata that provide a definitive description of what the data represents, and the spatial and temporal properties of the data. The OGC CF-netCDF standard consists of a suite of standards that support encoding of digital geospatial information representing space/time-varying phenomena. Although it was originally developed for the Earth science community, netCDF can be used to communicate and store a wide variety of multidimensional data. The netCDF data model and encodings are particularly well suited to providing data in forms familiar to atmospheric and oceanic scientists, specifically, as sets of related arrays.

Zarr Storage Specification (Zarr): Zarr [OGC 21-050r1] is an open-source specification for the storage of multidimensional arrays and associated metadata. Zarr stores metadata using .json text files and array data as (optionally) compressed binary chunks. The Zarr

specification details the contents and layout of these elements. Zarr can store data into any storage system that can be described as a key/value store. In a standard filesystem, the keys are filenames within a directory hierarchy, and the values are the file contents. In a cloud object store (e.g., Amazon S3), the keys are the object IDs and the values are the object data. This flexibility allows implementations to experiment with novel storage technologies while maintaining a uniform API for downstream libraries and users.

4.5 Metadata

Content standards may encode information in an interoperable manner, but those data are useless if they cannot be found. Metadata associated with the content data helps users understand the nature and usefulness of specific datasets. Metadata standards ensure that such descriptions are maintained in a common fashion so that users can search metadata catalogs and identify appropriate data for their requirements. Not only does metadata describe the contents, quality, usage rights, and source of the data, but properly built metadata allows for discoverability and assessment of the usefulness of the data which it describes. Geospatial metadata standards are maintained by ISO.

ISO 19115 Geospatial Information – Metadata: this is the most commonly used metadata standard for the geospatial community [ISO 19115]. The current version was approved in 2014 and has since been amended, but many organizations still use the 2003 edition, causing some degree of interoperability challenges in catalogs. To further complicate matters, many nations and communities have profiled the ISO 19115 standards to contain just a subset of information of particular interest to that organization. ISO 19115 contains additional parts to define metadata for gridded data as well as a schema.

ISO 19139 Geographic information — XML schema implementation: this standard [ISO 19139] complements ISO 19115 by providing an XML encoding ruleset for metadata. Such encoding are the most common means to store and exchange metadata records in the geospatial community. It is important to note that JSON is becoming an increasingly important data exchange mechanism and ISO 19115 JSON encoding is in work.