14

Geospatial Semantic Web

14.1 Introduction

The large volume of geospatial data available on the Web opens up unprecedented opportunities for data access and data interchange, facilitating the design of new geospatial applications and claiming for the redesign of traditional ones. However, to interoperate, geospatial applications must be able to locate and access the data sources, and agree both on the syntax and on the semantics of the data that flow between them. Achieving interoperability is more difficult in this case simply because geospatial data are much more complex than conventional data, with respect to both syntax and semantics.

A convenient way to provide universal access (over the Web) to geospatial data sources is to implement *geospatial Web services* that encapsulate the data sources, because this strategy obviates the need for traditional means of data distribution. Furthermore, as pointed out in Lieberman et al. (2005), geospatial Web services would benefit from the adoption of Semantic Web technologies, thereby becoming understandable to the applications. This evolution includes the development and encoding of formal geospatial ontologies, which would leverage existing standards. The result is called the *Geospatial Semantic Web*.

A major enabling technology of the Geospatial Semantic Web is the service-oriented architecture, which in turn depends on interoperability standards related to all aspects of geospatial service operations. The geospatial community has already developed a set of specifications through standard-setting bodies, including the *ISO Technical Committee 211* (TC 211), the *Open Geospatial Consortium* (OGC), the European Community initiative *Infrastructure for Spatial Information in Europe* (INSPIRE), and the U.S. *Federal Geospatial Data Committee* (FGDC).

The geomatics standardization activity under the ISO TC 211 is developing a family of international standards that will: (1) support the understanding and usage of geospatial information; (2) increase the availability, access, integration, and sharing of geospatial information, enabling interoperability of geospatial computer systems; and (3) facilitate the establishment of geospatial infrastructures on the local, regional, and global level. The work is divided into over 40 projects that resulted or will result in standard definitions.

The OGC comprises more than 310 companies, government agencies, and universities with the purpose of creating and promoting the development of technologies that facilitate interoperability between systems that process spatial and location data. The OGC working groups publish their results as interface specifications and standards for data interchange.

INSPIRE is an initiative of the European Commission to implement a European Spatial Data Infrastructure, built on the principle that it must be possible to seamlessly combine spatial information from different sources across Europe, and share it between many users and applications.

The FGDC is a U.S. interagency committee that coordinates the sharing of geospatial data through a portal that searches metadata held within the National Spatial Data Infrastructure (NSDI) Clearinghouse Network. The network is a collection of metadata catalogues, known as Clearinghouse Nodes. Each node is hosted by a different organization and contains metadata that describe geospatial data within the area of responsibility of the organization.

This chapter overviews the technologies that facilitate the development of the Geospatial Semantic Web, emphasizing the role of standard proposals. Each technology is first discussed from a broad perspective and then illustrated with implemented applications.

14.2 Basic Geospatial Concepts

In this section, we summarize a minimum set of geospatial concepts required for the understanding of the next sections.

Geospatial data relates to geography, location, addresses, or a place on the Earth’s surface. The term is often used in place of other terms such as maps, geographic data, or spatial data, but it also includes digital imagery of the Earth. Examples are property records, elevation, migratory routes, crime scenes, and burned areas.

Geospatial services are applications that process geospatial data. There are many types of geospatial services, such as Web mapping services, spatial analysis services, metadata services, and visualization services. A service may be tightly coupled with a specific dataset or dataset collection, or loosely coupled, otherwise.

A *feature* is an abstraction of a real world phenomenon and a *geospatial feature* is a feature associated with a location relative to the Earth (Percivall 2002). Geospatial features are usually classified into types (or classes), such as bays, rivers, cities, and so on, with well-defined characteristics. Feature types are in turn organized according to specific aspects, such as rivers and creeks are hydrographic feature types, and cities and environmental protection areas are administrative areas.

Note that geospatial datasets typically contain representations of the locations of a set of features, as in a map, or portray a set of features, as in a remote sensing image. We address the problem of classifying geospatial features in Section 14.3.

A *geospatial field* is a function that maps points of a domain region on the Earth's surface into values from a given set, the range of the function. For example, a vegetation map associates each point of the domain region with a specific type of vegetation, whereas a gravimetric map associates each point of the domain region with a real value.

A *coordinate reference system* (CRS) defines how geospatial data relate to real locations on the Earth's surface. The specific details of the CRS used for geospatial data need not concern us here. It suffices to point out that the specification of the location of a geospatial feature or the domain region of a geospatial field must include the CRS used.

A *gazetteer* (Fitzke and Atkinson 2006) defines a vocabulary consisting of identifiers, location descriptions, and attributes for a set of geospatial features, typically from a well-defined region, such as a country. A gazetteer therefore helps applications locate the desired geospatial features by their names or by a limited set of attribute values. We discuss gazetteers in Section 14.4.

A *metadata catalogue* stores descriptions of datasets housed in data sources (Nebert and Whiteside 2005), offering services to query and update metadata, and to request datasets from the data sources. Note that, typically, a catalogue does not store the datasets themselves. We describe two geospatial metadata standards in Section 14.5, and discuss geospatial metadata catalogues in Section 14.6.

As defined in Chapter 7, a *Web service* is a software system designed to support machine-to-machine interaction over a network. It has an interface described in WSDL and interacts with other components using SOAP messages. A *geospatial Web service* is a Web service that manipulates geospatial data. We briefly address geospatial Web services in Section 14.7.

Finally, a *spatial data infrastructure* is an integrated set of services that allow users to identify and access geospatial data from several sources. We provide two examples of spatial data infrastructures in Section 14.8.

14.3 Classifying Geospatial Features

In this section, we address how to define geospatial feature types (or classes). We first discuss how to construct *feature type thesauri*, which is the preferred way to organize geospatial features in a gazetteer. Then, we expand the discussion to consider ontologies that explore more complex relationships than those thesauri offer.

14.3.1 Geospatial Feature Type Thesauri

Recall from Chapter 2, that a *thesaurus* is defined as "the vocabulary of a controlled indexing language, formally organized so that a priori relationships between concepts (for example as 'broader' and 'narrower') are made explicit" (ISO 1986).

Table 14.1 Thesaurus relationships according to the ISO-2788 guidelines.

Relationship	Abbr.	Description
Scope Note	**SN**	a note attached to a term to indicate its meaning within an indexing language.
Use	**USE**	the term that follows the symbol is the preferred term when a choice between synonyms or quasi-synonyms exists.
Use For	**UF**	the term that follows the symbol is a nonpreferred synonym or quasi-synonym.
Top Term	**TT**	the term that follows the symbol is the name of the broadest class to which the specific concept belongs; sometimes used in the alphabetical section of a thesaurus.
Broader Term	**BT**	the term that follows the symbol represents a concept having a wider meaning.
Narrower Term	**NT**	the term that follows the symbol refers to a concept with a more specific meaning.
Related Term	**RT**	the term that follows the symbol is associated, but is not a synonym, a quasi-synonym, a broader term, or a narrower term.

A thesaurus usually provides a binary relationship between terms and annotations (Scope Note) and six binary relationships between terms, as summarized in Table 14.1.

The ADL Gazetteer effort exemplifies the difficulties of constructing a feature type thesaurus (Hill et al. 1999; Hill 2000). The ADL Feature Type Thesaurus (FTT) is the result of amalgamating a number of other gazetteer classifications and geospatial dictionaries. The FTT Version 2.0 of July, 2002 contains nearly 200 top terms and 3000 relationships. The topmost terms are:

```
1. administrative areas
2. hydrographic features
3. land parcels
4. manmade features
5. physiographic features
6. regions
```

Table 14.2 shows the *Narrow Term* hierarchy rooted at `regions`. Table 14.3 lists the same *Narrow Term* relationships, together with a few entries of the *Used For* and *Related Term* relationships involving such terms, to illustrate how thesauri relationships are used in the FTT.

Table 14.2 The narrow term/broader term hierarchy rooted at `regions`.

regions . agricultural regions . biogeographic regions . . barren lands . . deserts . . forests . . . petrified forests . . . rain forests . . . woods . . grasslands . . habitats . . jungles . . oases . . shrublands . . snow regions . . tundras . . wetlands	regions (cont.) . climatic regions . coastal zones . economic regions . land regions . . continents . . islands . . . archipelagos . . subcontinents . linguistic regions . map regions . . chart regions . . map quadrangle regions . . UTM zones

Table 14.3 Selected relationships involving terms under region.

Term	Rel.	Term
regions	NT	biogeographic regions
regions	NT	cadastral areas
regions	NT	climatic regions
regions	NT	coastal zones
regions	NT	economic regions
regions	NT	firebreaks
regions	NT	land regions
regions	NT	map regions
regions	NT	research areas
regions	RT	ocean regions
land regions	UF	lake districts
land regions	UF	lake regions
land regions	NT	continents
land regions	NT	islands
land regions	NT	subcontinents
land regions	BT	regions
islands	UF	atolls
islands	UF	cays
islands	UF	island arcs
islands	UF	isles
islands	UF	islets
islands	UF	keys (islands)
islands	UF	land-tied islands
islands	UF	mangrove islands
islands	NT	archipelagos
islands	RT	bars (physiographic)
islands	BT	land regions

islands » *A feature type category for places such as the island of Manhattan.*

Used for: » *The category **islands** is used instead of any of the following.*

- atolls
- cays
- island arcs
- isles
- islets
- keys (islands)
- land-tied islands
- mangrove islands

Broader Terms: land regions » ***islands** is a subtype of "land regions.."*
Related Terms: *» The following is a list of other categories related to **islands** (non-hierarchical relationships)*

- bars (physiographic)

Scope Note: Tracts of land smaller than a continent, surrounded by the water of an ocean, sea, lake or stream. [Glossary of Geology, 4th ed.]. *» Definition of **islands** .*

Fig. 14.1 Complete thesaurus entry for the term `islands` with explanations.

Figure 14.1 shows the complete FTT entry for the term `islands` with explanations where (sentences after the symbol "»" are treated as comments):

- The *Scope Note (SN)* clause shows the definition adopted for the term `islands`
- The *Used For (UF)* clauses indicate the nonpreferred terms that should be replaced by the term `islands`
- The *Broader Term (BT)* clause indicates that the term `islands` is classified under the term `land regions`
- The *Related Terms (RT)* clause points to other thesauri terms related to `islands`

In the construction of the ADL Feature Type Thesaurus, the choice of the preferred terms took into account their frequency of use in the reference sources. Closely related terms were entered as nonpreferred terms and linked to the preferred term via *Related Term* relationships, as illustrated in Fig. 14.1. The depth of the hierarchy was determined based on the specificity needed by the ADL gazetteer. Terms more specific than needed were entered as nonpreferred terms pointing to the broader term. For example, `atolls` are a specific type of `islands` that the FTT defines as a nonpreferred term related to `islands` by the *Used For* relationship. Therefore, a user searching for features classified as `atolls` will be redirected to `islands`.

14.3.2 Geospatial Feature Ontologies

As summarized at the end of Section 2.2.3, organizing a set of terms as a thesaurus has intrinsic limitations: (1) the set of term relationships is fixed, as shown in Table 14.1; and (2) term definitions are not formalized, in the sense that they are just natural language text linked to terms by Scope Notes.

Such limitations become apparent, for example, when one tries to automatically align two thesauri, even if they refer to the same concepts, for various reasons. The syntax of the terms is not a safe indication that they denote the same concept; the thesauri may in fact be written in different (natural) languages. The *Narrow Term* relationships (i.e., the hierarchy of terms) are likely to be different, reflecting the original purposes of the thesauri. One cannot directly rely on an automatic analysis of term definitions because they are not formalized. Lastly, a simple analysis of the definitions based on the frequency of occurrences of similar words might end up with wrong alignments.

Consider, for example, the definitions for the term `bay` from the ADL gazetteer and for the terms `bay` and `island` from the GEOnet Names Server (GNS) (URL: http://gnswww.nga.mil/geonames/GNS/index.jsp):

1. (ADL) bay: *indentation*s of a *coast*line or shoreline enclosing a part of a body of **water**; bodies of **water** partly **surrounded** by **land.**
2. (GNS) bay: a *coast*al *indentation* between two capes or headlands, larger than a cove but smaller than a gulf.
3. (GNS) island: tracts of **land**, smaller than a continent, **surrounded** by **water** at high **water.**

The reader may verify that the term `bay` from ADL has three words in common with the term `island` from GNS ("**water**", "**surrounded**", and "**land**"), and two with the term `bay` from GNS, even considering just the radix of the words ("*coast*al" "*indentation*s"). Therefore, an automatic alignment tool that compares term definitions based only on word frequency might end up aligning the term `bay` from ADL with the term `island` from GNS.

A strategy to circumvent this problem would be to use a geospatial vocabulary borrowed from an upper-level ontology, such as OpenCYC, as a *pivot vocabulary*. In this strategy, each thesaurus would be manually aligned with the pivot vocabulary. That is, each term from the thesaurus would be linked to a term of the pivot vocabulary, say, by a *Scope Note* relationship. Hence, two terms from different thesauri would be aligned if they were linked to the same term of the pivot vocabulary.

A much more sophisticated approach to create classifications for geospatial features would be to adopt full-fledged ontologies, which provide much more powerful constructs to define and relate classes. By doing so, one would be able to resort to OWL reasoners to analyze the classifications. In the rest of this subsection, we discuss three examples of such ontologies.

#$Island *islands*

A specialization of both #$LandBody and #$IslandOrIslandGroup. Each instance of #$Island is a body of land surrounded by water. #$Islands are typically much smaller in area than (instances of the similarly-defined collection of) #$TrueContinents (q.v.), though it would be rather arbitrary to try to distinguish these types on the basis of size alone. (And note that #$ContinentOfAustralia, e.g., is both an #$IslandOrIslandGroup (q.v.) and a #$TrueContinent.) For groups of #$Islands that form a geographical cluster, see #$Archipelago.

guid: bd58bb39-9c29-11b1-9dad-c379636f7270
direct instance of: #$ConventionalClassificationType #$ExistingObjectType
direct specialization of: #$LandBody #$IslandOrIslandGroup

(Cyc® Knowledge Base Copyright© 1995-2002 Cycorp, Inc., Austin, TX, USA.)

Fig. 14.2 Definition of the term `island` from OpenCYC.

The OpenCYC Selected Vocabulary and Upper Ontology — Geography (OpenCYC 2002) provides an example of a classification for geospatial features, using a proprietary ontology language. For example, Fig. 14.2 shows the definition of the term island.

Originally written in SUO-KIF, the Suggested Upper Merged Ontology (SUMO) (SUMO 2006; Niles and Pease 2001) organizes concept definitions into three levels (with the corresponding namespaces):

- SUMO Top Level: high-level abstractions of SUMO (http://reliant.teknowledge.com/DAML/SUMO.owl#)

- MILO (Mid-Level Ontology): a bridge between the abstract content of SUMO and the rich details of the various domain ontologies (http://reliant.teknowledge.com/DAML/Mid-level-ontology.owl#)

- Ontology of Geography: contains a number of classes that correspond to geospatial concepts, and instances that represent geospatial features; organized in the parts: (I) the Geography Terms for the CIA World Fact Book; and (II) the General Geography Terms and Background (http://reliant.teknowledge.com/DAML/Geography.owl)

Figure 14.3 contains a fragment of SUMO, in OWL, ending on the class `Island`. This fragment, albeit fairly long, does not come close to illustrating the richness of the definitions that SUMO contains.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:owl="http://www.w3.org/2002/07/owl#"
         xmlns="http://reliant.teknowledge.com/DAML/SUMO.owl#">

<!--  copyright 2003 (c) Teknowledge  -->
...

<rdfs:Class rdf:ID="Region">
  <rdfs:subClassOf rdf:resource="#Object"/>
  <rdfs:comment>A topographic location. Regions encompass surfaces
of Objects, imaginary places, and Geographic Areas. Note that a
Region is the only kind of Object which can be located at itself.
Note too that Region is not a subclass of SelfConnectedObject,
because some Regions, e.g. archipelagos, have parts which are not
connected with one another.
  </rdfs:comment>
</rdfs:Class>
...

<rdfs:Class rdf:ID="GeographicArea">
  <rdfs:subClassOf rdf:resource="#Region"/>
  <rdfs:comment>A geographic location, generally having definite
boundaries. Note that this differs from its immediate superclass
Region in that a GeographicArea is a three-dimensional Region of the
earth. Accordingly, all astronomical objects other than earth and
all one-dimensional and two-dimensional Regions are not classed
under GeographicArea.
  </rdfs:comment>
</rdfs:Class>
...

<rdfs:Class rdf:ID="LandArea">
  <rdfs:subClassOf rdf:resource="#GeographicArea"/>
  <rdfs:comment>An area which is predominantly solid ground, e.g. a
Nation, a mountain, a desert, etc. Note that a LandArea may contain
some relatively small WaterAreas. For example, Australia is a
LandArea even though it contains various rivers and lakes.
  </rdfs:comment>
</rdfs:Class>
...

<rdfs:Class rdf:ID="Island">
  <rdfs:subClassOf rdf:resource="#LandArea"/>
  <rdfs:comment>A LandArea that is completely surrounded by a
WaterArea.</rdfs:comment>
</rdfs:Class>
...
</rdf:RDF>
```

Fig. 14.3 A fragment of SUMO ontology ending on the class `Island`.

The *Semantic Web for Earth and Environmental Terminology* (SWEET) provides an upper-level ontology for the Earth system sciences (SWEET 2006; Raskin 2006). The SWEET ontologies include several thousand terms, grouped into facet ontologies, whose classes are connected by properties. For example, the *EarthRealm* ontology covers physical properties of the planet. It contains classes corresponding to realms, such as `Ocean`, and subclasses corresponding to the associated subrealms, such as “ocean floor”. For comparison, Figures 14.4a, 14.4b, and 14.4c show the definition of the class `island`.

```
<?xml version="1.0" encoding="UTF-8" ?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:owl="http://www.w3.org/2002/07/owl#"
 xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
 xmlns:substance="http://sweet.jpl.nasa.gov/ontology/substance.owl#"
 xmlns:space="http://sweet.jpl.nasa.gov/ontology/space.owl#"
 xmlns:numerics="http://sweet.jpl.nasa.gov/ontology/numerics.owl#"
 xmlns:property="http://sweet.jpl.nasa.gov/ontology/property.owl#"
 xmlns="http://sweet.jpl.nasa.gov/ontology/earthrealm.owl#">

<owl:Ontology rdf:about="">
  <dc:title>Earthrealm</dc:title>
  <dc:date>1/27/2004 2:59:58 PM</dc:date>
  <dc:creator>SWEET project</dc:creator>
  <dc:description></dc:description>
  <dc:subject></dc:subject>
  <owl:versionInfo>1.0</owl:versionInfo>
  <owl:imports
   rdf:resource="http://purl.org/dc/elements/1.1/"/>
  <owl:imports
  rdf:resource="http://sweet.jpl.nasa.gov/ontology/substance.owl" />
  <owl:imports
   rdf:resource="http://sweet.jpl.nasa.gov/ontology/space.owl" />
  <owl:imports
   rdf:resource="http://sweet.jpl.nasa.gov/ontology/numerics.owl" />
  <owl:imports
   rdf:resource="http://sweet.jpl.nasa.gov/ontology/property.owl" />
</owl:Ontology>
...
<owl:Class rdf:ID="Ocean">
  <rdfs:subClassOf rdf:resource="#BodyOfWater" />
  <rdfs:subClassOf>
    <owl:Class>
      <owl:complementOf>
        <owl:Class rdf:about="#LandWaterObject" />
    </owl:complementOf>
  </owl:Class>
</rdfs:subClassOf>
</owl:Class>
```

Fig. 14.4a A fragment of the SWEET Earthrealm Ontology ending on the class `Island`.

```
<owl:Class rdf:ID="LandRegion">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource=
"http://sweet.jpl.nasa.gov/ontology/space.owl#isPartOf" />
      <owl:allValuesFrom>
        <owl:Class>
          <owl:unionOf rdf:parseType="Collection">
            <owl:Class rdf:about="#LandSurfaceLayer" />
            <owl:Class rdf:about="#LandwaterSurfaceLayer" />
          </owl:unionOf></owl:Class>
      </owl:allValuesFrom>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasEcosystemType" />
      <owl:allValuesFrom rdf:resource="#TerrestrialEcosystem" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class>
      <owl:complementOf>
        <owl:Class rdf:about="#OceanRegion" />
      </owl:complementOf>
    </owl:Class>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="#TopographicalRegion" />
</owl:Class>
...
<owl:Class rdf:ID="OceanRegion">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource=
"http://sweet.jpl.nasa.gov/ontology/space.owl#isPartOf" />
      <owl:allValuesFrom rdf:resource="#Ocean" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasEcosystemType" />
      <owl:allValuesFrom rdf:resource="#MarineEcosystem" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="#TopographicalRegion" />
</owl:Class>
...
```

Fig. 14.4b A fragment of the SWEET Earthrealm Ontology ending on the class `Island`.

```
<owl:Class rdf:ID="LandwaterRegion">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource=
"http://sweet.jpl.nasa.gov/ontology/space.owl#isPartOf" />
      <owl:someValuesFrom rdf:resource="#LandwaterSurfaceLayer" />
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="#LandRegion" />
</owl:Class>
...
<owl:Class rdf:ID="Island">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource=
"http://sweet.jpl.nasa.gov/ontology/space.owl#surroundedBy_2D" />
      <owl:allValuesFrom>
        <owl:Class>
          <owl:unionOf rdf:parseType="Collection">
            <owl:Class rdf:about="#OceanRegion" />
            <owl:Class rdf:about="#LandwaterRegion" />
          </owl:unionOf>
        </owl:Class>
      </owl:allValuesFrom>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="#LandRegion" />
</owl:Class>
...
</rdf:RDF>
```

Fig. 14.4c A fragment of the SWEET Earthrealm Ontology ending on the class `Island`.

To summarize, OpenCYC and SUMO provide little more than a hierarchy of geospatial concepts, just as the ADL Feature Type Thesaurus does. The SWEET ontology goes further and uses OWL class restrictions to capture the semantics of the geospatial concepts. In turn, such class restrictions depend on the SWEET facets that model concepts related to time, space, and numerical extents (Raskin 2006).

In general, the creation of a rich geospatial feature ontology depends on defining specific ontologies to capture the semantics of geospatial metric operators and topological relationships, as already recognized in Egenhofer (2002) and in Hobbs (2005). However, as pointed out in Raskin (2006), a deficiency of RDF and OWL is that these languages provide no direct support for numerical concepts, and must rely on the XML Schema datatypes, which do not include operations or relations over those concepts.

Finally, as a benchmark for geospatial feature ontologies, one may select a small set of GNS (or ADL) definitions and check if they can be formalized in the ontology. For example, consider the GNS definition of “island” as “a tract of land, smaller than a continent, surrounded by water at high water.” To formalize this concept, the ontology must capture the notions of “tract”, “smaller”, “surrounded by”, and “high”, provided that the concepts of “land”, “continent”, and “water” are already in the ontology.

14.4 Gazetteers

14.4.1 Examples of Gazetteers

The common-sense geospatial reference system relies on explicitly naming geospatial features. This form of geospatial reference is usually supported by geospatial dictionaries, called *gazetteers*, containing lists of geospatial names, together with their geospatial locations and other descriptive information.

Most atlases have a gazetteer section that can be used to look up a geospatial name and find the pages where the corresponding geospatial feature appears. Digital gazetteers, readily available on the Web, have been developed to support geospatial information systems. They typically store the name, type, a point-based footprint, and a parent administrative area for each geospatial feature. They also provide various services to handle the stored data.

We now briefly describe some well-known digital gazetteers. The Getty Thesaurus of Geospatial Names (TGN 2004) is a structured, world-coverage vocabulary of 1.3 million names, coordinates, and other information for around 892,000 geospatial places. For each place name, the TGN maintains a unique id, a set of place types taken from the Art and Architecture Thesaurus (see Section 12.2.1), alternative versions of the name, a containing administrative region, and a footprint in the form of a point in latitude and longitude.

The U.S. Geological Survey's *Geospatial Names Information System* (GNIS) and the National Imagery and Mapping Agency's *Geospatial Names Processing System* (GNPS) are maintained in cooperation with the U.S. Board of Geospatial Names (BGN). GNIS contains information about almost 2 million physical and cultural geospatial features in the United States and its territories (GNIS 2005). The *GEOnet Names Server* (GNS) provides access to the National Geospatial-Intelligence Agency and the BGN database of foreign geospatial names, containing about 4 million features with 5.5 million names.

The Alexandria Digital Library (ADL) Project (Smith and Frew 1995; Smith 1996; Hill et al. 1999; Frew et al. 2000) is a research project for modeling, educational applications, and software components. The ADL Project also developed HTML clients to access the ADL collections and gazetteer.

The ADL Gazetteer prototype to evaluate digital library architectures, and gazetteer applications, has approximately 5.9 million geospatial names, classified according to the ADL Feature Type Thesaurus, discussed in Section 14.3.1. It combines data from the GNIS (U.S. names) and the GNPS (non-U.S. names). The ADL Catalogue stores the georeferenced holdings of the Map & Imagery Laboratory (UCSB Davidson Library), with more than 2 million items. Metadata for the catalogue holdings combine elements taken from the USMARC standard (USMARC 1976) and from the FGDC metadata scheme, discussed in Section 14.5.1. The complete metadata scheme has about 350 elements.

Finally, we mention Geonames (URL: http://www.geonames.org), which provides free access to a database with over six million entries for geographical names. The data are accessible through a user interface, as well as through Web services that offer geocoding based on geographical names or postal codes.

14.4.2 Standards for Gazetteers

In this section, we briefly discuss standards for gazetteer data and interfaces. The geomatics standardization activity under ISO Technical Committee 211 published a standard for gazetteer data, the ISO 19112:2003 *Spatial Referencing by Geospatial Identifiers* (ISO 2003), which defines a conceptual schema and guidelines for describing indirect spatial (noncoordinate) reference systems. Table 14.4 lists some of the attributes defined for gazetteer entries.

The ADL Gazetteer Content Standard supports rich descriptions of geospatial features that go beyond those of most traditional gazetteers. Table 14.5 contains the major sections of the ADL standard, which may be further divided into subsections for encoding more detailed information. For example, the streetAddressSection is divided into streetAddress, city, stateProvince, postalCode, and country.

Table 14.4 Selected attributes of the ISO 19112 conceptual schema.

Attribute	Comment
geospatial identifier	a unique identifier for the feature taken from a given identifier space
alternative identifiers	one or more alternative identifiers
name	the preferred name for the feature
type	the type of the feature, typically taken from a feature type thesaurus
description	a description of the feature
position	coordinates of a representative point of the feature
temporal extent	a description of the temporal extension of the feature
geospatial extent	a description of the geospatial extension of the feature
administrator	organization responsible for maintaining the feature’s name
parent	identifier of the parent feature of which the entry is a subdivision
child	identifiers of the features that the entry is a parent

Table 14.5 Selected attributes of the ADL Gazetteer Content Standard.

Attribute	Comment
featureID	unique identifier for the feature
featureName	the names for the feature
classSection	the primary type of the feature
codeSectionType	the code associated with the feature
spatialLocation	the map location of the feature
streetAddressSection	the street address of the feature
relatedFeature	relationships of the feature with other features
description	a short narrative description of the feature
featureData	data about the feature, such as population or elevation
featureLink	Web site that provides information on the feature
supplementalNote	note explaining unusual circumstance of the feature
entryMetadata	documents about entry and modification dates

The Open GIS Consortium has been developing a Web Gazetteer Service standard (Fitzke and Atkinson 2006) for distributed access to gazetteers. The gazetteer service interface specifies four operations that can be requested by a client and performed by a gazetteer service:

- *GetCapabilities* (required implementation by servers): allows a client to request and receive back service metadata (or *capabilities*), which are documents that describe the abilities of the specific server implementation. This operation also supports negotiation of the specification version being used for client-server interactions.

- *DescribeFeatureType* (mandatory implementation by servers): allows a client to retrieve an XML schema document that describes the structure of any feature type the server can service.

- *GetFeature* (mandatory implementation by servers): allows a client to retrieve feature instances. In addition, the client should be able to specify which feature properties to fetch and should be able to constrain the query spatially and non-spatially.

- *GetGMLObject* (optional implementation by servers): allows a client to retrieve element instances by traversing XLinks that refer to their XML IDs.

14.5 Geospatial Metadata

14.5.1 Geospatial Metadata Standards

Recall from Chapter 9 that the term *metadata* designates "data about other data," that is, metadata describe the data content, historic information about the data, how the data were obtained, and so on. Therefore, metadata facilitate interoperability by describing details about the data, which applications may explore to access the data. In fact, Haas and Carey (2003) point out that the lack of proper metadata is one of the top ten reasons why federated databases may fail.

We consider geospatial metadata schemas specified by two authorities, the U.S. Federal Geospatial Data Committee (FGDC) and the International Organization of Standards (ISO).

The FGDC defined the *Content Standard for Digital Geospatial Metadata* (CSDGM) (FGDC 1998), originally published in 1994 and revised in 1998. The FGDC is currently leading the development of a U.S. profile of the international metadata standard, ISO 19115:2003, discussed next.

The geomatics standardization activity under ISO Technical Committee 211 published, among others, three international metadata specifications:

- ISO 19115:2003 provides metadata elements about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geospatial data (ISO 2003)
- ISO 19119:2005 provides metadata elements to describe geospatial information services (ISO 2005)
- ISO 19139 defines a formal encoding and structure of ISO metadata for exchange (ISO 2006)

A catalogue service should therefore support metadata according to the ISO 19115:2003 and ISO 19119:2005, and publish XML encoding consistent with the ISO 19139 or profiles thereof. These and related standards can be found in the form of ontologies in Islam et al. (2004).

In the next three subsections, we briefly outline the CSDGM and the ISO 19115:2003 and ISO 19119:2005 specifications.

14.5.2 The FGDC Metadata Standard

The *Content Standard for Digital Geospatial Metadata* (CSDGM) (FGDC 1998) establishes a common set of terminologies and definitions for the documentation of digital geospatial data, including metadata elements for the following topics (FGDC 2000):

- *Identification*: name, developer, geospatial area covered, information themes included, access restrictions, currentness
- *Data Quality*: positional and attribute accuracy, completeness, consistency, provenance

- *Spatial Data Organization*: spatial data model used to encode the spatial data, number of spatial objects, methods other than coordinates, such as street addresses, used to encode locations
- *Spatial Reference*: encoding of coordinate locations, map projection or grid system used, parameters to convert the data to another coordinate system
- *Entity and Attribute Information*: geospatial information (roads, houses, elevation, temperature, etc.) included, information encoding
- *Distribution*: distribution agency, formats and media available, online availability, price
- *Metadata Reference*: timestamp and agency responsible for the metadata compilation

Table 14.6 lists the CSDGM essential metadata elements and Fig. 14.5 shows a sample XML metadata record with such elements.

Table 14.6 Required CSDGM XML tags*.

XML Tag / Metadata Element	Description
Data Originator /metadata/idinfo/citation/citeinfo/origin/	the name of an organization or individual that developed the dataset
Data Title /metadata/idinfo/citation/citeinfo/title/	the name by which the dataset is known
Abstract /metadata/idinfo/descript/abstract/	a brief narrative summary of the dataset
Progress /metadata/idinfo/status/progress	the state of the dataset
West Bounding Coordinate /metadata/idinfo/spdom/bounding/westbc/	westernmost coordinate of the limit of coverage expressed in longitude
East Bounding Coordinate /metadata/idinfo/spdom/bounding/eastbc/	easternmost coordinate of the limit of coverage expressed in longitude
North Bounding Coordinate /metadata/idinfo/spdom/bounding/northbc/	northernmost coordinate of the limit of coverage expressed in latitude
South Bounding Coordinate /metadata/idinfo/spdom/bounding/southbc/	southernmost coordinate of the limit of coverage expressed in latitude
Theme Keyword /metadata/idinfo/keywords/theme/themekey/	common-use word or phrase used to describe the subject of the dataset
Metadata Contact Organization /metadata/metainfo/metc/cntinfo/cntorgp/cntorg/	the name of the organization to which the contact type applies
Metadata Contact Person /metadata/metainfo/metc/cntinfo/cntperp/cntper/	the name of the individual to which the contact type applies
Metadata Contact Address City /metadata/metainfo/metac/cntinfo/cntaddr/city/	the city of the address

*Adapted from (FGDC 2006b).

Table 14.6 Required CSDGM XML tags (cont.)*.

XML Tag	Description
Metadata Element	
Metadata Contact Address State or Province	the state or province of the address
/metadata/metainfo/metac/cntinfo/cntaddr/state/	
Metadata Contact Address Postal Code	the ZIP or other postal code of the address
/metadata/metainfo/metac/cntinfo/cntaddr/postal/	
Publication Date	the date when the dataset is published or otherwise made available for release
/metadata/idinfo/citation/citeinfo/pubdate/	
Purpose	a summary of the intentions with which the dataset was developed
/metadata/idinfo/descript/purpose/	
Time Period of Content: Single Date	time period(s) for which the dataset corresponds to the currentness reference
/metadata/idinfo/timeperd/timeinfo/sngdate/caldate	
Time Period of Content: Range of Dates, Beginning Date	
/metadata/idinfo/timeperd/timeinfo/rngdates/begdate/	
Time Period of Content: Range of Dates, Ending Date	
/metadata/idinfo/timeperd/timeinfo/rngdates/enddate/	
Currentness Reference	the basis on which the time period of content information is determined
/metadata/idinfo/timeperd/current/	
Maintenance and Update Frequency	the frequency with which changes and additions are made to the dataset after the initial dataset is completed
/metadata/idinfo/status/update	
Theme Keyword Thesaurus	reference to a formally registered thesaurus or a similar authoritative source of theme keywords
/metadata/idinfo/keywords/theme/themekt/	
Access Constraints	restrictions and legal prerequisites for accessing the dataset
/metadata/idinfo/accconst/	
Use Constraints	restrictions and legal prerequisites for using the dataset after access is granted
/metadata/idinfo/useconst/	
Metadata Contact Address Type	the information provided by the address
/metadata/metainfo/metc/cntinfo/cntaddr/addrtype/	
Metadata Contact Phone number	the telephone number by which individuals can speak to the organization or individual
/metadata/metainfo/metc/cntinfo/cntvoice/	
Metadata Date	the date that the metadata were created or last updated
/metadata/metainfo/metd	

*Adapted from (FGDC 2006b).

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
   <metadata>
   <idinfo>
     <citation>
       <citeinfo>
           <origin>  Louisiana State University Coastal
                   Studies Institute
        </origin>
                <pubdate>20010907</pubdate>
                <title> Geomorphology and Processes of
                     Land Loss in Coastal Louisiana,
                     1932 - 1990
             </title>
       </citeinfo>
    </citation>
    <descript>
     <abstract>A raster GIS file that identifies the land loss
process and geomorphology associated with each 12.5 meter
pixel of land loss between 1932 and 1990. Land loss processes
are organized into a hierarchical classification system that
includes subclasses for erosion, submergence, direct removal,
and undetermined. Land loss geomorphology is organized into a
hierarchical classification system that includes subclasses
for both shoreline and interior loss.
     </abstract>
     <purpose>The objective of the study was to determine the
land loss geomorphologies associated with specific processes
of land loss in coastal Louisiana.
     </purpose>
    </descript>
    <timeperd>
      <timeinfo>
         <rngdates>
             <begdate>1932</begdate>
             <enddate>1990</enddate>
         </rngdates>
      </timeinfo>
      <current>ground condition</current>
    </timeperd>
    <status>
      <progress>Complete</progress>
      <update>None planned</update>
    </status>
    <spdom>
      <bounding>
         <westbc>-92.000057</westbc>
         <eastbc>-88.81416</eastbc>
         <northbc>30.498417</northbc>
         <southbc>28.914905</southbc>
      </bounding>
    </spdom>
   <keywords>
```

```
      <theme>
         <themekt>ISO 19115 Topic Category</themekt>
         <themekey>biota</themekey>
      </theme>
      <theme>
         <themekt>none</themekt>
         <themekey>land loss</themekey>
         <themekey>wetlands</themekey>
         <themekey>geomorphology</themekey>
         <themekey>landscape ecology</themekey>
      </theme>
   </keywords>
   <accconst>none</accconst>
   <useconst>The metadata should be read completely prior to
use of the dataset.
Data were collected and compiled as 12.5 meter pixels and
should not be extended beyond the reasonable limits of the
resolution. This is not a survey data product and should not
be utilized as such.
   </useconst>
   </idinfo>
   <metainfo>
      <metd>20010907</metd>
      <metc>
        <cntinfo>
          <cntorgp>
            <cntorg> Louisiana State University Coastal
                  Studies Institute
            </cntorg>
          </cntorgp>
          <cntaddr>
            <addrtype>mailing and physical address</addrtype>
            <city>Baton Rouge</city>
            <state>LA</state>
            <postal>70803</postal>
          </cntaddr>
          <cntvoice>(225) 578-2395</cntvoice>
        </cntinfo>
      </metc>
      <metstdn> FGDC Content Standards for Digital
             Geospatial Metadata
      </metstdn>
      <metstdv>FGDC-STD-001-1998</metstdv>
   </metainfo>
 </metadata>
```

Fig. 14.5 A sample XML metadata record with CSDGM essential elements (FGDC 2006b).

14.5.3 The ISO 19115:2003 Metadata Standard

The ISO 10115:2003 introduces an extensive set of metadata elements, of which, typically, only a subset is used. It presents metadata as UML packages, containing UML classes, which have attributes that identify the *metadata elements*. That is, a metadata element and a class attribute are synonymous. A data dictionary accompanies the UML description.

Tables 14.7a and 14.7b list the core metadata elements, selected to answer the following questions: "Does a dataset on a specific topic exist ('what')?", "For a specific place ('where')?", "For a specific date or period ('when')?", and "A point of contact to learn more about or order the dataset ('who')?"

The ISO 19115:2003 also defines a set of *topic categories*, shown in Table 14.8, which offer a high-level geospatial data thematic classification to assist in the grouping and search of available geospatial datasets.

Finally, to illustrate the use of the ISO 19115:2003, Fig. 14.6 shows the first few entries of a sample metadata, taken from Annex I of ISO (2003).

Table 14.7a ISO 19115:2003 Mandatory Core metadata elements[1,2].

Metadata Element	Description
Dataset title (Q)	name by which the cited resource is known
Dataset reference date (Q)	reference date for the cited resource
Dataset language	language(s) used within the dataset
Dataset topic category	main theme(s) of the dataset
Abstract describing the dataset (Q)	brief narrative summary of the content of the resource(s)
Metadata point of contact	identification of and means of communicating with person(s) and organization(s) using the resource(s)
Metadata date stamp	date that the metadata was created

Notes:
1. Tables 14.7a and 14.7b are adapted from Annex B — Data dictionary for geospatial metadata of ISO (2003).
2. The indication "Q" after the name of the elements is explained in Section 14.6.

Table 14.7b ISO 19115:2003 Conditional Core metadata elements.

Metadata Element	Description
Geographic location of the dataset (Q) (by 4 coordinates or by geographic id)	geographic area of the dataset
Dataset character set	full name of the character coding standard used for the dataset
Metadata language	language used for documenting metadata
Metadata character set	full name of the character coding standard used for the metadata set
Dataset responsible party	identification of, and means of communication with, persons and organizations associated with the dataset
Spatial resolution of the dataset	level of detail expressed as a scale factor or a ground distance
Distribution format (Q)	provides a description of the format of the data to be distributed
Additional extent information for the dataset (vertical or temporal)	information about horizontal, vertical, and temporal extent
Spatial representation type	digital mechanism used to represent spatial information
Reference system (Q)	description of the spatial and temporal reference systems used in the dataset
Lineage	general explanation of the data producer's knowledge about the lineage of a dataset
Online Resource	information about online sources from which the dataset, specification, or community profile name and extended metadata elements can be obtained
Metadata file identifier	unique identifier for this metadata file
Metadata standard name	name of the metadata standard (including profile name) used
Metadata standard version	version (profile) of the metadata standard used

Table 14.8 ISO 19115:2003 topic categories.

Name	Definition / Examples
	Definition
	Examples
farming	rearing of animals or cultivation of plants
	agriculture, irrigation, aquaculture, plantations, herding, pests/diseases affecting crops and livestock
biota	flora and/or fauna in natural environment
	wildlife, vegetation, biological sciences, ecology, wilderness, sea life, wetlands, habitat
boundaries	legal land descriptions
	political and administrative boundaries
climatologyMeteorologyAtmosphere	processes and phenomena of the atmosphere
	cloud cover, weather, climate, atmospheric conditions, climate change, precipitation
economy	economic activities, conditions and employment
	production, labor, revenue, commerce, industry, tourism and ecotourism, forestry, fisheries, commercial or subsistence hunting, exploration and exploitation of resources such as minerals, oil, and gas
elevation	height above or below sea level
	altitude, bathymetry, digital elevation models, slope, derived products
environment	environmental resources, protection, and conservation
	environmental pollution, waste storage and treatment, environmental impact assessment, monitoring environmental risk, nature reserves, landscape
geoscientificInformation	information pertaining to earth sciences
	geophysical features and processes, geology, minerals, sciences dealing with the composition, structure, and origin of the earth’s rocks, risks of earthquakes, volcanic activity, landslides, gravity information, soils, permafrost, hydrogeology, erosion
health	health, health services, human ecology, and safety
	disease and illness, factors affecting health, hygiene, substance abuse, mental and physical health, health services

Table 14.8 ISO 19115:2003 topic categories (cont.).

Name	Definition / Examples
imageryBaseMapsEarthCover	base maps
	land cover, topographic maps, imagery, unclassified images, annotations
intelligenceMilitary	military bases, structures, activities
	barracks, training grounds, military transportation, information collection
inlandWaters	inland water features, drainage systems, and their characteristics
	rivers and glaciers, salt lakes, water utilization plans, dams, currents, floods, water quality, hydrographic charts
location	positional information and services
	addresses, geodetic networks, control points, postal zones and services, place names
oceans	features and characteristics of salt water bodies (excluding inland waters)
	tides, tidal waves, coastal information, reefs
planningCadastre	information used for appropriate actions for future use of the land
	land use maps, zoning maps, cadastral surveys, land ownership
sociology	characteristics of society and cultures
	settlements, anthropology, archaeology, education, traditional beliefs, manners, and customs, demographic data, recreational areas and activities, social impact assessments, crime and justice, census information
structure	man-made construction
	buildings, museums, churches, factories, housing, monuments, shops, towers
transportation	means and aids for conveying persons and/or goods
	roads, airports/airstrips, shipping routes, tunnels, nautical charts, vehicle or vessel location, aeronautical charts, railways
utilitiesCommunication	energy, water, and waste systems and communications infrastructure and services
	hydroelectricity, geothermal, solar and nuclear sources of energy, water purification and distribution, sewage collection and disposal, electricity and gas distribution, data communication, telecommunication, radio, communication networks

MD_Metadata
+identificationInfo
MD_DataIdentification
citation:
. CI_Citation
. title: Exploration Licences for Minerals
. date:
. CI_Date
. dateType: 001
. date: 193001
abstract: Location of all current mineral Exploration Licences issued under the Mining Act, 1971. Exploration Licences provide exclusive tenure rights to explore for mineral resources for up to a maximum of 5 years. Comment is sought on applications for Exploration Licences from numerous sources before granting. Exploration programs are subject to strict environmental and heritage conditions. Exploitation of identified resources must be made under separate mineral production leases.
...
extent:
.. EX_Extent
…….. +geographicElement
…….. EX_GeographicBoundingBox
.. …. westBoundLongitude: 129.0
.. …. eastBoundLongitude: 141.0
.. …. southBoundLatitude: -26.0
.. …. northBoundLatitude: -38.5
…….. description: South Australia
language: en
...
+referenceSystemInfo
MD_ReferenceSystem
referenceSystemIdentifier:
. RS_Identifier
. code: GDA 94
. codeSpace: DIPR
fileIdentifier: ANZSA1000001233
language: en
characterSet: 001
contact:
CI_ResponsibleParty
role: 002
organisationName: Department of Primary Industries and Resources SA
dateStamp: 20000803
metadataStandardName: ISO 19115
metadataStandardVersion: FDIS
dataset: https://info.pir.sa.gov.au/geometa/migs/MIGS_Down_cat.jsp

Fig 14.6 Sample metadata following the ISO 19115:2003 standard (Annex I of ISO (2003)).

14.5.4 The ISO 19119:2005 Service Metadata Standard

The ISO 10119:2005 introduces metadata elements that describe services, rather than data (Percivall 2002; ISO 2005). The metadata elements for a service provide sufficient information to allow a client to invoke the service based on the metadata record. If a service instance is tightly coupled with a dataset instance, the service metadata describes both the service and the geospatial dataset, using the ISO 19115 specification for the latter. For the loosely coupled case, dataset metadata need not be provided in the service metadata.

Table 14.9 summarizes the mandatory service metadata elements, and Table 14.10 introduces a geospatial services taxonomy. Table 14.11 indicates how the ISO 19100 series standards relate to the services taxonomy.

Table 14.9 ISO 19119:2005 mandatory service metadata elements*.

Section	Element	Description
Service Identification		provides a general description of the service sufficient to allow a client to invoke the service
	serviceType	a service type name from a registry of services
	containsOperations	provides information about the operations that comprise the service
Operation Metadata		describes the signature of one and only one method provided by the service
	operationName	a unique identifier for this interface
	DCP	distributed computing platforms on which the operation has been implemented
	connectPoint	handle for accessing the service interface
Service Provider		describes an organization that provides services
	providerName	a unique identifier for this organization
	serviceContact	information for contacting the service provider
Data Identification		(data available from a particular service, defined using elements of ISO 19115)
Operation Chain Metadata	name	the name, as used by the service for this chain
Parameter	name	the name, as used by the service for this parameter
	Optionality	indication if the parameter is required
	Repeatability	indication if more than one value of the parameter may be provided

*Adapted from Annex C, data dictionary for geospatial service metadata (ISO 2005).

Table 14.10 Geospatial services taxonomy.

Classification	Selected Examples
Geographic human interaction services	- Catalogue viewer - Geographic viewer - Geographic feature editor - Geographic symbol editor
Geographic model/ information management services	- Feature access service: provides a client access to and management of a feature store - Map access service: provides a client access to pictures of - geographic data - Catalogue service: provides discovery and management services on a store of metadata about instances - Gazetteer service: provides access to a directory of instances of a class or classes of real-world phenomena containing some information regarding position
Geographic workflow/ task management services	- Chain definition service: service to define a chain and to enable it to be executed by the workflow enactment service - Workflow enactment service: interprets a chain and controls the instantiation of services and sequencing of activities
Geographic processing services	- modify feature attributes
spatial	- Coordinate conversion service - Dimension measurement service
thematic	- Thematic classification service - Spatial counting service
temporal	- Temporal reference system transformation service - Temporal proximity analysis service
metadata	- Geographic annotation services
Geographic communication services	- Encoding service - Transfer service
Geographic system management services	(none identified).

Table 14.11 Mapping ISO 19100 series standards to service categories.

Service Category	Relevant ISO 19100 Series Standard
Geographic human interaction services	19117 Geographic information — Portrayal
	19128 Geographic information — Web Map server interface
Geographic model/ Information management services	19107 Geographic information — Spatial schema
	19110 Geographic information — Methodology for feature cataloguing
	19111 Geographic information — Spatial referencing by coordinates
	19112 Geographic information — Spatial referencing by geographic identifiers
	19115 Geographic information — Metadata
	19123 Geographic information — Schema for coverage geometry and functions
	19125-1 Geographic information — Simple feature access — Part 1: Common architecture
	19128 Geographic information — Web Map server interface.
Geographic Workflow/ Task management services	(no relevant ISO 19100 series standards)
Geographic processing service	19107 Geographic information — Spatial schema
	19108 Geographic Information — Temporal schema
	19109 Geographic information — Rules for application schema
	19111 Geographic information — Spatial referencing by coordinates
	19116 Geographic information — Positioning services
	19123 Geographic information — Schema for coverage geometry and functions
	19118 Geographic information — Encoding
Geographic communication services	(no relevant ISO 19100 series standards)
Geographic system management services	(no relevant ISO 19100 series standards)

14.6 The OGC Catalogue Specification

The OGC Catalogue Services Specification Catalogue (OGC)Catalogue (Nebert and Whiteside 2005) is a general model for catalogue services. It defines interfaces to publish and access digital catalogues of metadata for geospatial data, services, and related resource information. The specification is in fact a framework, in the sense that it leaves open, to a certain extent, the specific catalogue scheme, the query language, the format of the result sets, and how the catalogue should be updated. This point is important to accommodate different catalogue implementations and increase the opportunities for interoperability.

In this section, we outline a specialization of the framework for managing metadata resources that comply with ISO 19115:2003 and ISO 19119:2005 (Senkler et al. 2004; Voges and Senkler 2006). The discussion that follows refers to this particular specialization.

The catalogue services support:

- *Publishing* metadata, that is, inserting new metadata into the catalogue, either by the owner of the georesource or by a broker that acts on behalf of the owner.

- *Discovering* metadata, either by browsing the content of the catalogue or querying the catalogue.

- *Harvesting* metadata, that is, uploading new metadata from records located elsewhere.

Figure 14.7 shows the reference architecture for OGC-compliant catalogue services. The architecture is a multitier arrangement of clients and servers. The application client interfaces with the catalogue service using the OGC catalogue interface. The catalogue service may respond to an application request by accessing a local metadata repository, a resource service, or another catalogue service.

The catalogue discovery interface offers the following operations:

- *GetCapabilities:* allows clients to retrieve service metadata from a server.

- *GetRecords:* allows clients to search the catalogue and receive the results of the search.

- *GetRecordById:* retrieves the default representation of catalogue records.

- *DescribeRecord:* allows a client to discover elements of the information model supported by the target catalogue service.

- *GetDomain:* used to obtain runtime information about the range of values of a metadata record element or request parameter.

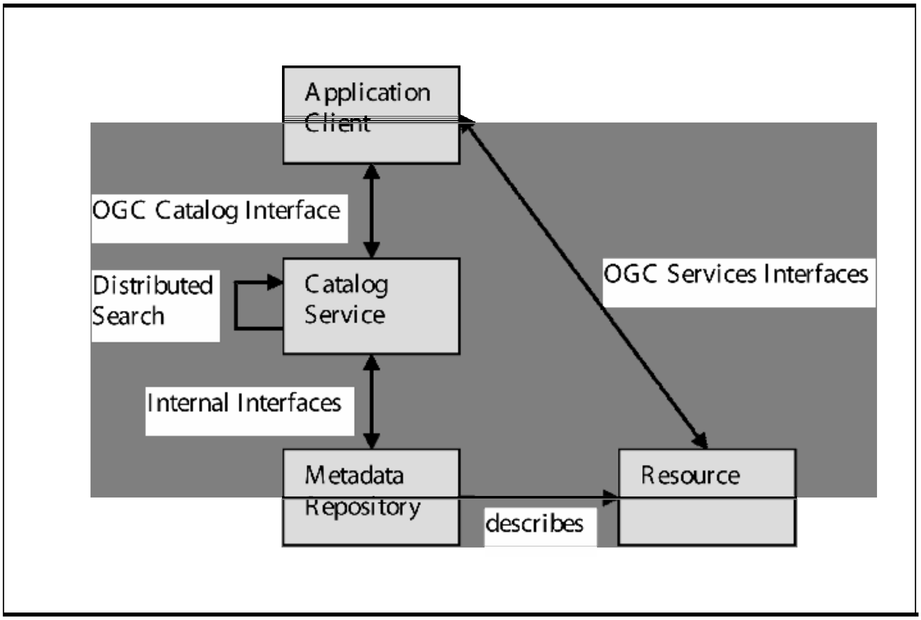

Fig. 14.7 The OGC Catalogue architecture.

The catalogue manager interface offers two operations:

- *Transaction Operation:* this operation allows clients to create, modify, and delete catalogue records (i.e., to "push" metadata into the catalogue).

- *Harvest Operation:* this operation only references the data to be inserted or updated in the catalogue, and it is the job of the catalogue service to resolve the reference, fetch those data, and process them into the catalogue (i.e., to "pull" metadata into the catalogue).

The specification defines a minimum set of *queryable* metadata elements, marked as "Q" in Table 14.7, to which *GetRecords* operations may refer (specific implementations may support additional elements). The specification also defines that the results must be returned as XML documents created using a new group of public metadata elements, expressed using the nomenclature and syntax of Dublin Core metadata. In other words, the metadata repository may have its own organization, as long as the catalogue services that encapsulate the repository accept this minimum set of queryable metadata elements and return results in the appropriate markup. This strategy, therefore, allows for maximum flexibility while guaranteeing minimum interoperability capability.

We leave an example of a catalogue federation to Section 14.8.2, where we discuss the U.S. National Spatial Data Infrastructure (NSDI).

14.7 Geospatial Web Services

In this section, we outline the effort of the Open Geospatial Consortium to define a set of specifications for geospatial Web services. The discussion is meant to instigate the reader to visit the OGC Web site (URL: www.opengeospatial.org) and browse through the documentation. Indeed, several OGC specifications have already become ISO standards, or permeated to SQL:2003, the latest version of the relational query language SQL (Eisenberg et al. 2004).

As an implementation example, we briefly address the *HDF-EOS Web GIS Software Suite (NWGISS) Project* (Di et al. 2001). Again, we refer the reader to the OGC Web site for a comprehensive list of implementations.

14.7.1 The OGC Services Framework

The *Open Geospatial Consortium (OGC)* specifications follow a reference architecture, called the *OpenGIS® Services Framework (OSF)* (Percivall 2002; Buehler 2003), that identifies services, interfaces, and exchange protocols that the applications may use. Briefly, the OpenGIS Services Framework consists of (see Fig. 14.8):

- *Coding Standards:* standards that specify geospatial data interchange and storage formats, including spatial reference systems and geospatial object geometry and topology. They include the *Geography Markup Language (GML).*

- *Application Services:* components that, on the client side, interact with the users and, on the server side, interact with the application and data servers.

- *Registry Services:* components that offer mechanisms to classify, register, describe, search, maintain, and access data about the resources available. They include the *Web Registry Service (WRS).*

- *Portrayal Services:* components that offer specific support for the visualization of geospatial data, in the form of maps, 3-D terrain perspectives, annotated images, and so on. They include the *Web Map Service (WMS)*, the *Coverage Portrayal Service (CPS)*, and the *Style Management Service (SMS)*.

- *Data Services:* components that offer basic access services to geospatial data. They include the *Web Object Service* (WOS), the *Web Feature Service* (WFS), the *Sensor Collection Service* (SCS), the *Image Archive Service* (IAS), and the *Web Coverage Service* (WCS).

- *Processing Services:* components that operate on geospatial data and provide value-added services for applications. They include gazetteer and catalogue services.

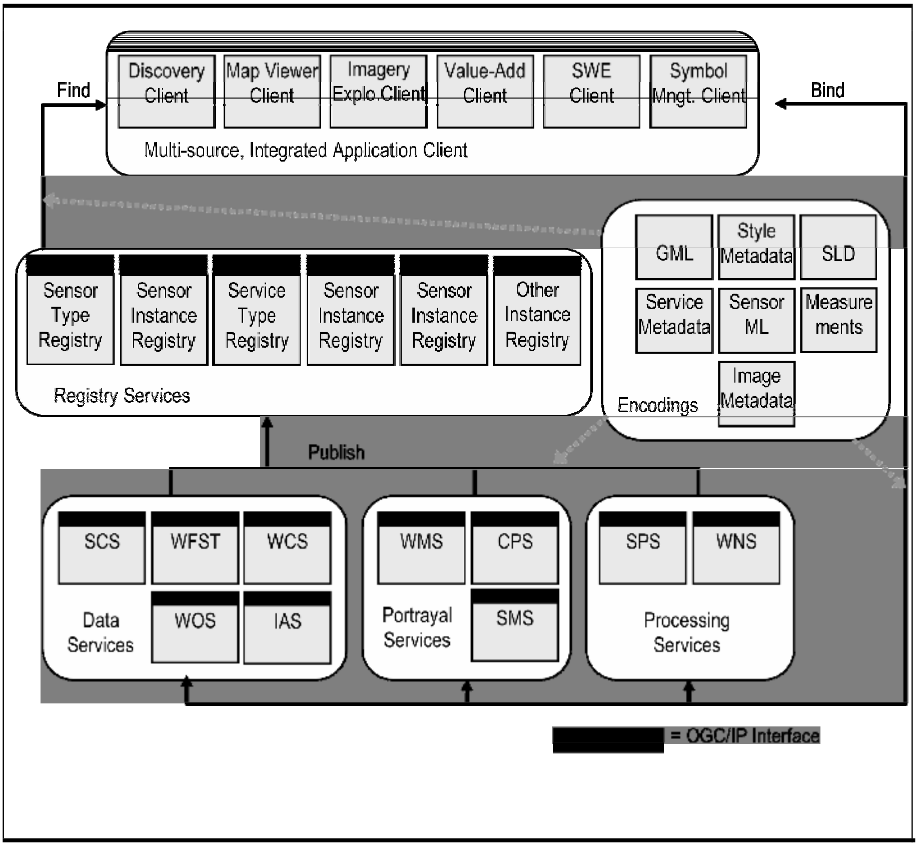

Fig. 14.8 Components of the OGC Web Service Framework (Percivall 2002).

The Geography Markup Language (GML) is an XML markup specification for geospatial data interchange (Cox et al. 2003), defined as an XML schema that contains models for geometries, features, surfaces, and the like. Applications may then define their own XML schemas based on the GML schema types.

The OGC services offer fairly uniform interfaces. For example, the Web Map Service (WMS) defines three interfaces:

- *GetCapabilities* permits a client application to obtain information about the WMS service and about the data stored in the WMS server.

- *GetMap* permits a client application to request a map (sent as an image) that portrays datasets stored in the WMS server.

- *GetFeatureInfo* permits a client application to request data associated with objects portrayed in the current map.

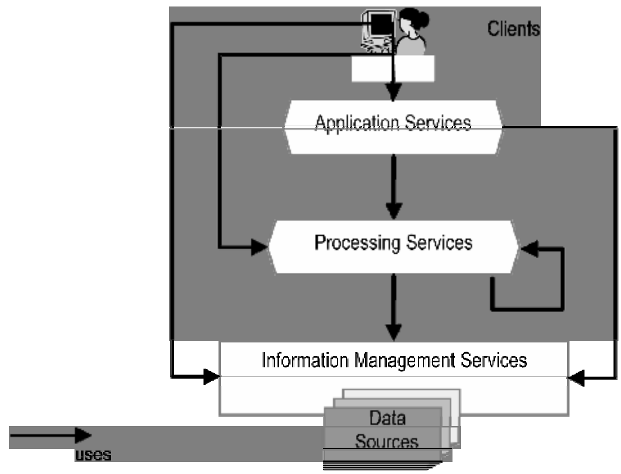

Fig. 14.9 Service tiers in the OWS architecture (Whiteside 2005).

14.7.2 A Geospatial Web Services Architecture

The OSF was developed independently of the W3C protocol architecture for Web services. Only recently, a series of specification proposals, collectively known as the *OpenGIS Web Service 2 Initiative* (Sonnet 2004; Whiteside 2005), defined interfaces that abide by the W3C standards. OGC also started a large-scale experiment to test the concept of a Geospatial Semantic Web (OGC 2005).

The *OGC Web Services (OWS) Architecture* is a service-oriented architecture, with all components providing one or more services to other services or to clients. This service-oriented architecture is based on the fundamental roles of service provider and service consumer within a distributed computing system. It also focuses on interactions among components implementing defined services, in the form of service requests, service responses, and service exceptions.

Following Whiteside (2005), we define:

- A *service* is a distinct part of the functionality that is provided by an entity through interfaces.
- An *interface* is a named set of operations that characterize the behavior of an entity.
- An *operation* is a specification of a transformation or query that an object may be called to execute. It has a name and lists of input and output parameters.

The OGC Web service interfaces use open standards and provide only a few static operations per service. Standard XML-based data encoding formats and languages are used in many server-to-client and client-to-server data transfers. Services are loosely organized into four tiers, as shown in Fig. 14.9.

The information management services tier contains adapters that encapsulate the data sources. The services at this tier usually include some processing of the data retrieved. For example, WMS, WCS, and WFS can perform coordinate transformation and format conversion. The processing services tier contains services designed to process data. The application services tier contains services designed to support clients, especially thin clients, such as Web browsers.

The OWS architecture supports *service chaining*, either *transparently* (defined and controlled by the client), *translucently* (predefined but visible to the client), or *opaquely* (predefined and not visible to client), according to ISO (2005).

14.7.3 Example of an Integrated Collection of Geospatial Web Services

Mission to Planet Earth

The *Mission to Planet Earth* (MTPE) is a NASA project designed to study our planet using data collected by a number of satellites. The *EOS Data and Information System (EOSDIS)* (Kobler et al. 1995) is the component responsible for providing access to the data generated in the context of the MTPE project. It is a distributed system with the following components.

- Distributed Active Archive Centers (DAACs): responsible for processing, archiving, and distributing EOS and related data, and for providing user support.

- Science Data Processing Segment (SDPS): responsible for supporting product generation, data archiving and distribution, and information management.

- Science Investigator-led Processing Systems (SIPSs): process EOS data into standard products.

- Networks, EOS Data and Operations System (EDOS): dedicated resource that supports inter-DAAC data flows for generation of interdependent EOS products.

- Flight Operations Segment (FOS): provides mission planning and scheduling, and monitors health and safety of the spacecraft and instruments.

At present, EOSDIS is managing and distributing data from: EOS missions (Landsat-7, QuikSCAT, Terra and, ACRIMSAT), Pre-EOS missions (UARS, SeaWIFS, TOMS-EP, TOPEX/Poseidon, and TRMM), and all of the Earth Science Enterprise legacy data (e.g., pathfinder datasets). EOSDIS handles extraordinary rates and volumes of scientific data. For example, the Terra spacecraft alone produces 194 GB of data per day, and downlinks data at 150 Mb/s.

Finally, HDF-EOS is the standard file format for storing EOS data.

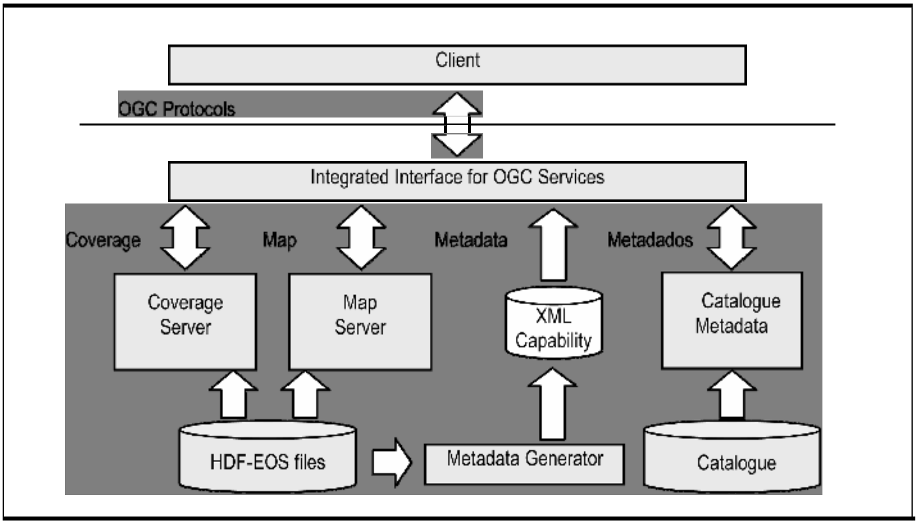

Fig. 14.10 HDF-EOS Web GIS software suite (NWGISS) architecture.

HDF-EOS Web GIS Software Suite (NWGISS) Project

The goal of the *HDF-EOS Web GIS Software Suite (NWGISS) Project* (Di et al. 2001) is to make EOS data available to other applications that follow the OGC specifications. The NWGISS architecture (see Fig. 14.10) consists of three major components: a map server, a coverage server, and a catalogue server. The NWGISS also includes an OGC WCS client.

The map server implements the OGC WMS services for all HDF-EOS data formats. In particular, it generates georeferencing information for a map at runtime, if necessary. The coverage server implements the OGC WCS services for three data formats: HDF-EOS, NITF, and GeoTIFF. The server may resample, cut, and reassemble coverages (geographic fields) in realtime, as well as apply format transformations.

The coverage client implements the OGC WCS specifications. The client acts as a mediator (Gupta et al. 2000) to access coverages in the HDF-EOS, NITF, and GeoTIFF formats, stored at any OGC WCS server (not just those the NWGISS implements). The client supports data access and visualization, among other operations.

The NWGISS client is a limited mediator that allows the user to select coverages from one or more sources, using their metadata, and to visualize them together, among other operations.

14.8 Examples of Spatial Data Infrastructures

A *spatial data infrastructure* (SDI) delivers integrated spatial information services, which should allow users to identify and access spatial or geographical information from a wide range of sources, from the local level to the global level, in an interoperable way for a variety of uses. One of the main elements of these infrastructures is a *geoportal*, which facilitates access to the spatial data and provides complementary services. *Catalogue geoportals* facilitate accessing metadata catalogues, whereas *application portals* provide dynamic geospatial Web services.

This section introduces two initiatives, the European Community INSPIRE, and the U.S. National Spatial Data Infrastructure (NSDI), the latter supported by a catalogue geoportal. Comprehensive surveys of SDI initiatives around the globe can be found in Crompvoets et al. (2005) and Bregt and Crompvoets (2005).

14.8.1 INSPIRE

Overview

The *Infrastructure for Spatial Information in Europe* (INSPIRE) is an initiative of the European Commission to implement a European Spatial Data Infrastructure, built on the basis of the following principles (INSPIRE 2004):

- Data should be collected once and maintained at the level where this can be done most effectively.
- It must be possible to seamlessly combine geospatial data from different sources across Europe and share it among many users and applications.
- It must be possible for data collected at one level to be shared among all the different levels, for example, detailed for detailed investigations, and general for strategic purposes.
- Geospatial data needed for good governance at all levels should be abundant and widely available under conditions that do not restrain its extensive use.
- It must be easy to discover which geospatial data are available, fit the needs for a particular use, and under what conditions they can be acquired and used.
- Geospatial data must become easy to understand and interpret because they can be visualized within the appropriate context and selected in a user-friendly way.

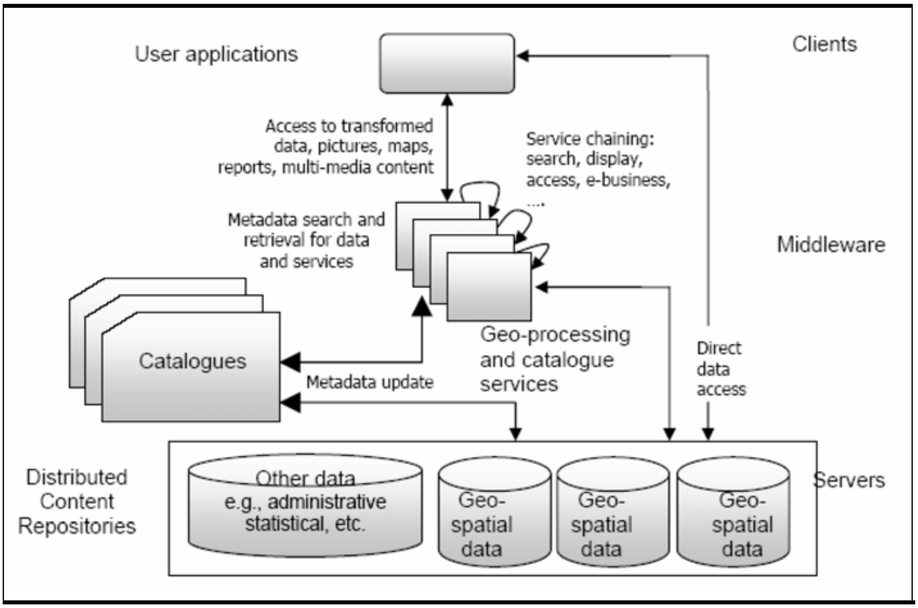

Fig. 14.11 The architecture reference model for INSPIRE.

Architecture

The INSPIRE architecture (Smits 2002) considers a network of SDIs that follow the reference model shown in Fig 14.11. The reference model distinguishes four major groups of components: user applications, geoprocessing and catalogue services, catalogues, and content repositories

User applications include general-purpose interfaces for querying, and viewing, a tool for database administrators, and analytical applications tailored to the information needs of the user. Geoprocessing and catalogue services may process user queries, draw maps from data, regulate access, perform payment operations, and extract and send data to a user application. Content repositories provide data. Finally, catalogues allow clients and services to find out what repositories or services are available and appropriate for their use.

Reference Data and Metadata

The INSPIRE Reference Data and Metadata (RDM) (Rase et al. 2002) describe the necessary geographical reference data and their metadata. The following components form the geographical reference data:

- Geodetic reference data
- Units of administration
- Units of property rights (parcels, buildings)
- Addresses

- Selected topographic themes (hydrography, transport, height)
- Orthoimagery
- Geographical names

The list of the common aspects include: geodetic reference system, quality, maintenance, interoperability, resolution/scale, and implementation priorities, language and culture, and metadata. We refer the reader to Rase et al. (2002) for a complete discussion of all such aspects.

As for metadata, the RDM group recommends that:

- All the reference data should be documented by metadata
- The three aspects of metadata must be considered: discovery, access, and use
- A metadata profile compatible with ISO 19115 must be developed
- Metadata shall be kept up to date
- The member states shall identify a competent authority for coordinating the national producers of data and for managing the metadata information systems
- Priority should be given to create a one-stop geoportal for discovering and accessing geospatial data

14.8.2 NSDI

Overview

The U.S. *National Spatial Data Infrastructure* (NSDI) is a virtual network designed to enable the development and sharing of U.S. digital geographic information resources. The goals of the NSDI are: to reduce duplication of efforts among agencies; to improve quality and reduce costs related to geographic information; to make geographic data more accessible to the public; to increase the benefits of using available data; and to establish key partnerships with states, counties, cities, tribal nations, academia, and the private sector to increase data availability.

The Federal Geographic Data Committee (FGDC) is charged with coordinating the development of the NSDI through three major activities:

- Establishment of a National Geospatial Data Clearinghouse
- Development of standards for data documentation, collection, and exchange
- Development of policies, procedures and partnerships to create a national digital geospatial data framework

NSDI Clearinghouse Network

The National Geospatial Data Clearinghouse Network (FGDC 2006c) is a community of distributed data providers who publish collections of metadata that describe their map and data resources within their areas of responsibility, documenting data quality, characteristics, and accessibility. Each metadata collection, known as a *Clearinghouse Node*, is hosted by an organization to publicize the availability of data within the NSDI. The metadata in these nodes are searched by the `geodata.gov` portal, discussed next.

The geodata.gov Portal

The `geodata.gov` is a catalogue of geospatial information containing metadata records and links to maps, features, catalogue services, downloadable datasets, images, clearinghouses, map files, and more. The metadata records are submitted to the portal directly by government agencies, individuals, and companies, or by harvesting the data from geospatial clearinghouses. The portal is operated in support of the *Geospatial One-Stop Initiative* to provide one-stop access to all registered geographic information, and related online access services within the United States.

The Geospatial One-Stop Initiative follows standards intended to establish common requirements to facilitate data exchange for seven themes of geospatial data that are of critical importance to the NSDI. The seven geospatial data themes are: geodetic control, elevation, orthoimagery, hydrography, transportation, cadastral, and governmental unit boundaries. These themes are known as *NSDI framework themes*. The Framework Data Content exchange standard is being processed as an American National Standard through the ANSI-accredited organization, INCITS. It is anticipated to become a formal standard in late 2006. Finally, the portal supports both the FGDC and the ISO metadata standards, discussed earlier in this chapter.

14.9 Example of a Metadata Catalogue for Earth Science Data

In this section, we overview the Global Change Master Directory (GCMD), a metadata catalogue for Earth science data, developed at the NASA Goddard Space Flight Center (Meaux 2005; Olsen 2005). The information in this section was extracted from the GCDM Web site (URL: http://gcmd.gsfc.nasa.gov) with permission (Olsen 2006).

14.9.1 Overview of the GCMD

The *Global Change Master Directory* (GCMD) is a metadata catalogue based on the DIF, Directory Interchange Format (DIF 2006), and the SERF, Services Entry Resources Format (SERF 2006). The DIF is used to create directory entries that describe a group of data. A DIF entry consists of a collection of fields that detail specific information about a group of data. DIF fields are compatible with the ISO 19115. SERF is a *de facto* standard used to create directory entries that describe a group of services. GCMD also supports a set of controlled science keywords.

GCMD allows a user to access subsets by organization (e.g., NOAA metadata records). In general, GCMD supports searches that include spatial and temporal restrictions, as well as restrictions on attributes such as data center, geographical location, science-data platform, topics, and so on. GCMD also offers a metadata authoring tool for adding information to the catalogue, which uses templates that minimize user interaction and contribute to correctness of the metadata.

Fig. 14.12 The Global Change Master Directory (GCMD) Web site.

As illustrated in Fig. 14.12, GCMD supports datasets for the following:

- Agriculture (forest science, soils, etc.)
- Atmosphere (precipitation, air quality, etc.)
- Biosphere (vegetation, zoology, etc.)
- Climate indicators (air temperature, drought conditions, etc.)
- Cryosphere (frozen ground, sea ice, etc.)
- Human dimensions (land use, populations, etc.)
- Hydrosphere (rivers/streams, water quality, etc.)
- Land surface (erosion, topography, etc.)
- Oceans (marine biology, salinity, etc.)
- Paleoclimate (icecores, land records, etc.)
- Solid Earth (geochemistry, seismology, etc.)
- Spectral/engineering (radar, visible imagery, etc.)
- Sun-Earth Interactions (auroras, solar activity, etc.)
- Data centers (locations, instruments, projects, platforms/sources)

The GCMD system supplies data services that can be found by topic, including:

- Data analysis and visualization (GIS, image processing, etc.)
- Data management/handling (archiving, reformatting, sub/supersetting, etc.)
- Education/outreach (interactive programs, etc.)
- Environmental advisories (fire, health, weather, climate, etc.)
- Hazards management (recovery/relief, mitigation, response, etc.)
- Metadata handling (authoring tools, discovery, transformation, conversion, etc.)
- Models (numerical, physical, etc.)
- Reference and information service (bibliographic, subscription services, etc.)

It is worth noting that GCMD is the American Coordinating Node of the Committee on Earth Observation Satellites (CEOS) International Directory Network. CEOS membership encompasses the world's government agencies responsible for civil Earth Observation (EO) satellite programs, along with agencies that receive and process data remotely acquired from space.

GCMD has recognized the importance of customization for partner organizations and is doing so by generating subset views of the GCMD catalogue through portals. Portals have made it easier for organizations to maintain and document their data in one place without duplicating the effort to create another online catalogue.

14.9.2 GCMD and Other Data Sources

GCMD is a mature project. Its value has been recognized both nationally and internationally. In what follows, we briefly discuss GCMD's collaborations with NASA, other U.S. government agencies, and international organizations as an indication of GCMD's maturity and value.

NASA Collaborations

GCMD/NASA collaborations include:

- *NASA Earth Observing System Data Information System (EOSDIS):* GCMD provides dataset descriptions for all of NASA's EOSDIS publicly available data holdings.

- *Federation of Earth Science Information Partners (ESIP):* GCMD collaborates directly with two ESIP and indirectly with others. ESIP include the Distributed Oceanographic Data System (DODS) and the Seasonal and Inter-annual Information Partnership (SIESIP). GCMD has provided a customized ESIP data portal and ESIP services portal to search for ESIP data.

U.S. Government Federal Agency Collaborations

GCMD/Federal Agencies collaborations include:

- *Biological Resources Division (BRD), U.S. Geological Survey:* The BRD provides the scientific understanding and technologies needed to support the sound management and conservation of U.S. biological resources. NASA and the BRD are in a partnership to increase access to ecological data through shared metadata population efforts of the GCMD and the National Biological Information Infrastructure.

- *Geospatial One-Stop (GOS):* Metadata records about NASA's earth science data are being contributed from GCMD to the GOS portal (see Section 14.8.2). GCMD provides access to project-level information as part of NASA's contribution to GOS. A custom portal is available to search NASA projects.

- *National Oceanic and Atmospheric Administration (NOAA):* GCMD continues close collaborations with NOAA through the exchange of metadata information. GCMD is a participant in the NOAA Operational Model Archive and Distribution System (NOMADS) program and has contributed a custom portal for model output datasets.

- *National Spatial Data Infrastructure/Federal Geographic Data Committee (NSDI/FGDC):* GCMD is an NSDI/FGDC participant and a node of the FDGC Clearinghouse (see Chapter 14.8.2). Geospatial data can be searched across many different spatial servers by accessing the clearinghouse gateway.

- *Data Management and Communications (DMAC):* GCMD has contributed to the U.S. Integrated Ocean Observing Program (IOOS) proposed Data Management and Communications (DMAC) system. The GCMD will participate in the search and discovery of IOOS datasets.

- *Marine Metadata Interoperability (MMI):* GCMD is collaborating with the MMI project, a community effort aimed at making marine science datasets easier to find, access, and use. Scientists and data managers can find valuable information about data management and metadata policies on the MMI Web site. GCMD contributes by sharing its extensive list of keywords and metadata expertise with the community.

International Collaborations

GCMD maintains several international collaborations, which include:

- *Committee on Earth Observation Satellites International Directory Network (CEOS IDN):* The CEOS IDN is an international effort to assist researchers in locating information on available datasets. GCMD shares information with other CEOS IDN nodes throughout the world and also provides software, search interfaces, and metadata writing tools.

- *Antarctic Master Directory:* GCMD is collaborating with the Joint Committee on Antarctic Data Management (JCADM), an international effort to make available Antarctic research data holdings. GCMD has developed an Antarctic Master Directory (AMD) to search for Antarctic datasets.

- *United Nations Environmental Programme (UNEP):* GCMD is continuing collaborations with the United Nations Environmental Programme (UNEP) Global Resources Information Database (GRID) to make available dataset descriptions from many nations throughout the world. The UNEP/GRID site in Budapest, Hungary, is a partner in the testing of new GCMD software technologies.

- *The Global Observing Systems Information Center (GOSIC):* GCMD is collaborating with the Global Observing Systems Information Center (GOSIC) to develop a portal to GOSIC data that will be collected by three major observing programs: the Global Climate Observing System (GCOS), the Global Ocean Observing System (GOOS), and the Global Terrestrial Observing System (GTOS). GCMD has contributed a customized Global Observation of Forest Cover (GOFC) portal GOFC, originally a joint program of the Committee on Earth Observation Satellites (CEOS) and GTOS, which now is part of the GTOS program.

- *The International Oceanographic Data and Information Exchange (IODE):* GCMD has collaborated with the International Oceanographic Data and Information Exchange (IODE) by testing the MEDI metadata tool. MEDI is compatible with the Directory Interchange Format (DIF) and IODE has encouraged their data centers to create new MEDI records and send metadata to the GCMD.

Additional examples of international GCMD collaborations are: the Global Ocean Ecosystems Dynamics Project (GLOBEC), the Gulf of Maine Ocean Data Partnership (GoMODP), and the Ocean Biogeographic Information System (OBIS).

Recommended Reading

In this chapter, we covered some of the more stable technologies that will lead to the Geospatial Semantic Web. However, we left out other useful technologies, such as geocoding, defined as "the process of linking words, terms and codes found in a text string to their applicable geospatial features, with known locations" (Margoulies 2001). We have also not touched on the design of search engines that deal with the spatial context implicit in Web documents (Hiramatsu and Reitsma 2004; Souza et al. 2005).

As a starting point for additional reading, we recommend the proceedings of the First International Conference on GeoSpatial Semantics (Rodríguez et al. 2005), especially the keynote address that discusses the development of a common ontology for expressing and reasoning about spatial information for the Semantic Web (Hobbs 2005).

References

ADL (2006) The Alexandria Digital Library Project — Gazetteer. Available at: http://www.alexandria.ucsb.edu/gazetteer/.

ADL (2002) ADL Gazetteer Content Standard. Available at: http://www.alexandria.ucsb.edu/gazetteer/FeatureTypes/ver070302/index.htm.

Bregt, A.; Crompvoets, J. (2005) Spatial data infrastructures: Hype or hit? In: *Proceedings From Pharaohs to Geoinformatics. FIG Working Week 2005 and GSDI*-8. Cairo, Egypt, April 16–21.

Buehler, K. (Ed.) (2003) OpenGIS® Reference Model. Document number OGC 03-040 Version: 0.1.3. Open Geospatial Consortium, Inc.

Cox, S.; Daisey, P.; Lake, R.; Portele, C.; Whitside, A. (Eds.) (2003). OpenGIS® Geography Markup Language (GML) Implementation Specification. Open Geospatial Consortium, Inc.

Crompvoets, J.; Bregt, A.; de Bree, F.; van Oort, P.; Lownen, B.; Rajabifard, A.; Williamson, I. (2005) Worldwide (Status, Development and) Impact Assessment of Geoportals. In: *Proceedings From Pharaohs to Geoinformatics. FIG Working Week 2005 and GSDI-8*. Cairo, Egypt, April 16–21.

Di, L.; Yang, W.; Deng, M.; Deng, M.; McDonald, K. (2001) The prototype NASA HDF-EOS Web GIS Software Suite (NWGISS). In: *Proceedings of the NASA Earth Science Technologies Conference*. Greenbelt, Maryland. August 28-30, (CD-ROM, 4pp). Available at: http://laits.gmu.edu/Papers/NWGISS.htm.

DIF (2006) *Directory Interchange Format (DIF) Writer's Guide*. Global Change Master Directory. National Aeronautics and Space Administration. Available at: http://gcmd.nasa.gov/User/difguide/.

Egenhofer, M.J. (2002) Toward the semantic geospatial Web. In: *Proceedings of the ACM GIS'02*, November 8–9, McLean, Virginia, USA.

Eisenberg, A.; Melton, J.; Kulkarni, K.G.; Michels, J-E..; Zemke, F. (2004) SQL: 2003 has been published. *ACM SIGMOD Record*, Vol. 33, No. 1, pp. 119–126.

FGDC (1998) Content Standard for Digital Geospatial Metadata. Metadata Ad Hoc Working Group, Federal Geographic Data Committee. FGDC-STD-001-1998. Available at: http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/v2_0698.pdf.

FGDC (2000) Content Standard for Digital Geospatial Metadata Workbook. Federal Geographic Data Committee. Available at: http://www.fgdc.gov/metadata/documents/workbook_0501_bmk.pdf.

FGDC (2005) Towards A National Geospatial Strategy & Implementation Plan. Federal Geographic Data Committee. Available at: http://www.fgdc.goc/futuredirections/.

FGDC (2006a) Geospatial Metadata Standards. Federal Geographic Data Committee. Available at: http://www.fgdc.gov/metadata/geospatial-metadata-standards.

FGDC (2006b) Creating and Publishing Metadata in Support of Geospatial One-Stop and the NSDI. Federal Geographic Data Committee. Available at: http://www.fgdc.gov/dataandservices/pub_guidance.

FGDC (2006c) NSDI Framework Introduction and Guide (Handbook). Federal Geographic Data Committee. Available at: http://www.fgdc.gov/framework/handbook/index_html.

Fitzke, J.; Atkinson, R. (2006) Gazetteer Service Profile of the Web Feature Service Implementation Specification, Version 0.9.1. OGC® Implementation Specification, OGC 05-035r1. Open Geospatial Consortium Inc.

Frew, J.; Freeston, M.; Freitas, N.; Hill, L.; Janée, G.; Lovette, K.; Nideffer, R.; Smith, T.; Zheng, Q. (2000) The Alexandria Digital Library architecture. *International Journal on Digital Libraries*, Vol. 2, No.4, pp. 259–268.

Fu, G.; Abdelmoty, A.; Jones, C. (2003) Design of a Geographical Ontology. SPIRIT - Spatially-Aware Information Retrieval on the Internet, EU IST Programme, D5 3101.

Gardels, K. (1996) The Open GIS approach to distributed geodata and geoprocessing. In: *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling.* Santa Fe, NM, USA, pp. 21–25.

GNIS (2005) Geographic Names Information System. U.S. Department of the Interior, U.S. Geological Survey, Reston, USA. Available at: http://geonames.usgs.gov/index.html.

Gupta, A.; Zaslavsky, I.; Marciano, R. (2000) Generating query evaluation plans within a spatial mediation framework. In: *Proceedings of the Ninth International Symposium on Spatial Data Handling*. Beijing, China, pp. 8a18–8a31.

Haas, L.; Carey, M. (2003) Will federated databases ever go anywhere? In: *Lowell Database Research Self-Assessment Meeting*. Lowell, Massachusetts, USA.

Hill, L.L. (2000) Core elements of digital gazetteers: Placenames, categories, and footprints. In: J. Borbinha and T. Baker (Eds.) *Research and Advanced Technology for Digital Libraries*, pp. 280–290.

Hill, L.L.; Frew, J.; Zheng, Q. (1999) Geographic names: The implementation of a gazetteer in a georeferenced digital library. *D-Lib*, January 1999.

Hiramatsu, K.; Reitsma, F. (2004) GeoReferencing the Semantic Web: Ontology-based markup of geographically referenced information. In: *Proceedings of the Joint EuroSDR /EuroGeographics Workshop on Ontologies and Schema Translation Services*. Paris, France.

Hobbs, J. (2005) Interoperability among geospatial ontologies (Keynote Address). In: Rodríguez, M.A.; Cruz, I.F.; Levashkin, S.; Egenhofer, M.J. (Eds.) (2005) *Proceedings of the First International Conference on GeoSpatial Semantics - GeoS 2005*, Mexico City, Mexico, November 29–30, 2005. Available at: http://colab.cim3.net/file/work/SICoP/2006-06-20/JHobbs06202006.ppt.

Hobbs, J.; Pan, F. (2006) Time Ontology in OWL. Ontology Engineering Patterns Task Force of the Semantic Web Best Practices and Deployment Working Group, World Wide Web Consortium (W3C) notes, Editor's Draft.

INSPIRE (2004) INSPIRE: INfrastructure for SPatial InfoRmation in Europe. Commission of the European Communities. Brussels, Belgium. Available at: http://inspire.jrc.it/proposal/EN.pdf.

ISO (1986) Documentation - Guidelines for the development of monolingual thesauri. International Standard ISO-2788, Second edition.

ISO (2003) Geographic information – Metadata. Draft International Standard ISO/FDIS 19115:2003(E). International Organization for Standardization.

ISO (2003b) Geographic information – Spatial Referencing by Geographic Identifiers. ISO International Standard 19112:2003.

ISO (2005) Geographic Information - Services. ISO 19119:2005. International Organization for Standardization.

ISO (2006) Geographic information – Metadata – Implementation specification. ISO 19139 (2006, v1.0).

Islam, A.S.; Bermudez, L.; Beran, B.; Fellah, S.; Piasecki, M. (2004) Ontologies for ISO Geographic Information Standards. Available at: http://loki.cae.drexel.edu/~wbs/ontology/2004/09/bug/iso-19115/index.htm.

Kobler, B.; Berbert, J.; Caulk, P.; Hariharan, P.C. (1995) Architecture and design of storage and data management for the NASA earth observing system data and information system (EOSDIS). In: *Proceedings of the Fourteenth IEEE Symposium on Mass Storage Systems*. Monterey, California, USA, p. 65.

Lieberman, J.; Pehle, T.; Dean, M. (2005) Semantic evolution of geospatial Web services. In: *Proceedings of the Workshop on Frameworks for Semantics in Web Services, DERI*, Innsbruck, Austria. June 9–10.

Maguire, D.J.; Longley, P.A. (2005) The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, Vol. 29, pp. 3–14.

Margoulies, S. (Ed.) (2001) Geocoder Service Specification. Version 0.7.6. OGC-IP Draft Candidate Implementation Specification OGC-01-026r1, Discussion Paper. Open Geospatial Consortium. Available at: http://feature.opengis.org/members/archive/arch01/01-026r1.pdf.

Meaux, M. (2005) Online metadata directories: A way of preserving, sharing and discovering scientific information. In: *Proceedings of the Gulf of Maine Ocean Data Partnership (GoMODP) Metadata Training Workshop*, Gulf of Maine Research Institute, Portland, ME, 19–20 October.

Mena, E.; Illarramendi, A.; Kashyap, V.; Sheth, A. (2000) OBSERVER: An approach for query processing in global information systems based on interoperation across pre-existing ontologies. *International Journal on Distributed and Parallel Databases (DAPD)*, Vol. 8, No. 2, pp. 223–272.

Nebert, D.; Whiteside, A. (Eds.) (2005) OpenGIS® Catalogue Services Specification, Version 2.0.0 with Corrigendum, OpenGIS® Implementation Specification OGC 04-021r3, Open Geospatial Consortium.

Niles, I.; Pease, A. (2001) Towards a standard upper ontology. In: C. Welty and B. Smith (Eds.). *Proceedings of the Second International Conference on Formal Ontology in Information Systems (FOIS-2001)*, Ogunquit, Maine, October 17–19, 2001.

Nogueras-Iso, J. (2005) Web ontology service, a key component of a spatial data infrastructure. In: *Proceedings of the Eleventh EC-GI & GIS Workshop, ESDI: Setting the Framework*. Alghero, Sardinia, 29th June – 1st July.

Nogueras-Iso, J.; Zarazaga-Soria, F.; Muro-Medrano, P. (2005) *Geographic Information Metadata for Spatial Data Infrastructures Resources, Interoperability and Information Retrieval.* New York, Springer-Verlag.

NSDI (2006) National Spatial Data Infrastructure. Available at: http://www.fgdc.gov/nsdi /nsdi.html.

OGC (2005) OGC to Begin Geospatial Semantic Web Interoperability Experiment. Press Release, 12/04/2005. Available at: http://www.opengeospatial.org/press/?page= pressrelease&year=0&prid=222.

Olsen, L. (2005) Data discovery worldwide using the Global Change Master Directory. In: *Proceedings of the Computing in Atmospheric Sciences Conference (CAS2K5)*, Annecy, France, 11–15 September.

Olsen, L. (2006) Private communication.

OpenCYC (2002) OpenCyc Selected Vocabulary and Upper Ontology - Geography. Cycorp. http://www.cyc.com/cycdoc/vocab/geography-vocab.html.

Percivall, G. (Ed.) (2002) The OpenGIS® Abstract Specification - Topic 12: OpenGIS® Service Architecture Version 4.3. OpenGIS® Project Document Number 02-112. Open GIS Consortium.

Rase, D.; Björnsson, A.; Probert, M.; Haupt, M-F. (Eds.) (2002) Reference Data and Metadata Position Paper. INSPIRE RDM PP v4-3 en. European Commission, Joint Research Centre. Available at: http://inspire.jrc.it/reports/position_papers/inspire_rdm_pp_v4_3_ en.pdf.

Raskin, R. (2006) Guide to SWEET Ontologies. NASA/Jet Propulsion Lab, Pasadena, CA, USA. Available at: http://sweet.jpl.nasa.gov/guide.doc.

Rodríguez, M.A.; Cruz, I.F.; Levashkin, S.; Egenhofer, M.J. (Eds.) (2005) *Proceedings of the First International Conference on GeoSpatial Semantics - GeoS 2005*, Mexico City, Mexico, November 29–30, 2005. Lecture Notes in Computer Science, Volume 3799, 2005, Berlin / Heidelberg, Springer.

Senkler, K.; Voges, U.; Remke, A. (2004) An ISO 19115/19119 Profile for the OGC Catalogue Services CSW 2.0. In: *Proceedings of the tenth EC GI & GIS Workshop, ESDI State of the Art*, Warsaw, Poland, 23–25 June.

SERF (2006) Service Entry Resource Format (SERF) Writer's Guide. Global Change Master Directory. National Aeronautics and Space Administration. Available at: http://gcmd.nasa.gov/serfguide/.

Smith, T.R. (1996) A digital library for geographically referenced materials. *IEEE Computer*, Vol. 29, No. 5, pp. 54–60.

Smith, T.R.; Frew, J. (1995) Alexandria Digital Library. *Communications of the ACM*, Vol. 38, No. 4, pp. 61–62.

Smits, P (Ed) (2002) INSPIRE Architecture and Standards Position Paper. INSPIRE AST PP v4-3 en. European Commission, Joint Research Centre. Available at: http://inspire.jrc.it/ reports/position_papers/inspire_ast_pp_v4_3_en.pdf.

Sonnet, J. (Ed) (2004) OWS 2 Common Architecture: WSDL SOAP UDDI, Version: 1.0.0. Discussion Paper OGC 04-060r1. Open Geospatial Consortium, Inc.

Souza, L.A. et al. (2005) The role of gazetteers in geographic knowledge discovery on the Web. In: *Proceedings of the Third Latin American Web Congress*. Buenos Aires, Argentina.

Stubkjær, E. (2001) Integrating ontologies: Assessing the use of the Cyc ontology for cadastral applications. In: *Proceedings of the Eighth Scandinavian Research Conference on Geographical Information Science - ScanGIS'2001*. Ås, Norway, 25th–27th June.

SUMO (2006) Suggested Upper Merged Ontology (SUMO). Available at: http://www.ontologyportal.org/.

SWEET (2006) The Semantic Web for Earth and Environmental Terminology (SWEET). Jet Propulsion Laboratory, California Institute of Technology. Available at: http://sweet.jpl.nasa.gov/index.html.

TGN (2004) Thesaurus of Geographical Names. The Getty Foundation. Available at: http://www.getty.edu/research/tools/vocabulary/tgn.

USMARC (1976) A MARC Format, in OFFICE, L. O. C. M. D., ed., Washington, D.C., Library of Congress Information Systems Office.

Voges, U.; Senkler, K. (2006) OpenGIS® Catalogue Services Specification 2.0.1 (with Corrigendum) - ISO Metadata Application Profile. OpenGIS® Application Profile OGC 04-038r4, Version: 1.0. Open Geospatial Consortium Inc.

Whiteside, A. (Ed) (2005) OpenGIS® Web services architecture description, Version: 0.1.0. OpenGIS® Best Practices, OGC 05-042r2. Open Geospatial Consortium Inc.

Wordnet (2005). Wordnet - A lexical database for the English language. Cognitive Science Laboratory, Princeton University, Princeton, NJ, USA. Available at: http://wordnet.princeton.edu.