Theme 6. OPEN SPATIAL ACCESS IN SDI

6.1. Preconditions of understanding of spatial access in SDI

1. Once spatial data of interest have been located and evaluated, using the Catalogue and Web Mapping techniques, access to detailed spatial data in its packaged form is often required by advanced users or application software.

2. Data access is recognized element in a full-service SDI.

6.2. Context and rationale of open spatial access in SDI

1. Access to spatial data from the consumers point of view is a part of a process of that goes from discovery to evaluation, to access and finally to exploitation:

1) **Discovery** (find, locate) involves the use of services such as metadata catalogues to find data of particular interest over a specific geographic region;

2) **Evaluation** involves detailed reports, sample data and visualization to help the consumer determine whether the data is of interest;

3) **Access** involves the order, packaging and delivery, offline or online, of the data (coordinate and attributes according to the form of the data) specified.

2. With the growth of the Internet access has become a demand driven operation.

3. Consumers expect simple discover and access to cheap (or free) data in simple standard formats that can be used in desktop applications.

4. The further democratization of access to geospatial data thus enables value-added suppliers to create new data products and services.

5. Several **trends** can be noted **in the treatment and handling of spatial data**:

1) Typically in the past **the first trend** concern of a data custodian has been what format the data is stored or managed in;

2) **Second trend** is in the organization of the data itself and a supplier was empowered to deliver data online;

3) More recently the third trend has been to merge all the discrete data sets together into single, seamless data warehouses that have spawned the development of direct data access services.

6.3. Organisational approach to open spatial access in SDI

6.3.1. Stakeholders' categories in SDI access

1. In most national infrastructures **government suppliers** are key stakeholders:

– How they will play in the development and operation of the data access component of the infrastructure depends strongly on government policies regarding data distribution, cost recovery, etc.

2. **Commercial entities** will generally play a strong role as providers of tools and services but may also be suppliers of primary and value-added data

3. The final category of stakeholder is **the consumer** or **end-user**:

– Their use of the data access element infrastructure is dependent on a number of factors including:

1) The functionality of the infrastructure tools;

2) The amount and quality of the content accessible;

3) Operating policies;

4) Infrastructure business model (will consumers be charged for access?), etc.

6.3.2. Development of policy/organisational environment for SDI access

1. Potential stakeholders will only become active participants if:

   1) They see advantages for their organizations;

   2) They do not feel threatened by the infrastructure.

2. This policy/organization environment:

   1) Will vary from country to country;

   2) Will need to be worked out closely with the stakeholder community.

3. The buy-in and commitment from senior management of

4. Some of the **issues that need to be considered in the development of the supportive to SDI access policy/organisational environment** are:

1) Distributed/autonomous suppliers;

2) The management of the data should be done as close as possible to source:

– This ensures the accuracy and quality of the data;

3) Commercial and government stakeholders need to feel comfortable as active participants in the infrastructure:

– They should not feel threatened by infrastructure business models or policies;

4) Multiple levels of “buy-in”; low barrier to entry:

a) The access component of the infrastructure must provide multiple levels of buying from a low cost option with limited benefits to higher cost options that offer increased benefits;

b) This allows suppliers to choose a level of participation that best meets their business and operational objectives;

5) Sustainable long term business models:

a) The access component of an infrastructure must provide an environment that supports a variety of supplier business models;

b) The development of a sustainable business model for the operation of the access component is critical to the long term success of the entire infrastructure.

5. **The role the private sector as suppliers of data, services, and technology and as potential operators of the SDI access** must be clearly defined.

6.3.3. Marketing and promotion of SDI access

1. The access component of an infrastructure must develop a **marketing and promotion plan** to build up the level of awareness and participation as quickly as possible.

2. It is important to get a critical mass of suppliers so potential participants will see the benefits of joining the infrastructure.

3. **Potential benefits to suppliers** include:

1) Economies of data collection, closest to the source;

2) Reduced operational costs;

3) New clients (national and international);

4) Data reuse (reuse vs. recollection or conversion);

5) Common tool and service reuse;

6) Advertising;

7) Benefits of 'free' portrayal;

8) Enabling/supporting broad new applications.

6.4. Implementation approach to open spatial access in SDI: definitions and overview

6.4.1. Data sets in SDI access

1. Data sets are described by metadata and maintained within a data store.

2. Core and Framework data sets are data that may be present within a spatial data infrastructure (See Theme 2).

3. Data sets are composed of:

1) Collections of features (e.g. roads, rivers, political boundaries, etc.);

2) Coverages (e.g. satellite/airborne imagery, digital

6.4.2. Data stores in SDI access

1. Data stores are used to manage data sets.

2. Data stores may be offline or online repositories.

3. Traditional online data stores are file-based repositories, setup for the delivery of pre-defined data sets.

4. Data stores also contain text and attribute data related to a data set.

5. **Data warehouses** are data stores that provide seamless access and management of data sets.

6.4.3 Spatial data warehouse in SDI access

1. A **spatial data warehouse** provides storage, management and direct access mechanisms.

2. **Key characteristics of a spatial data warehouse** include:

1) The access and delivery of arbitrary features, layers, etc.;

2) Seamless repository;

3) Common data model;

4) Application neutral, supporting a heterogeneous application environment;

5) Support of large volumes of data;

6) Multi-temporal support;

7) Common repository for spatial and non-spatial data;

3. Examples of commercial data warehousing and service solutions for geospatial data include:

a) Cubestore from Cubewerx (*http://www.cubewerx.com/*);

b) The Oracle Spatial solution (*http://otn.oracle.com/products/oracle9i/datasheets/spatial/spatial.html*);

c) ESRI Spatial Data Engine (*http://www.esri.com/*).

6.4.4. Data access service in SDI

1. Implementations of data access services include the following **services**:

1) Offline service (e.g. packaging and physical delivery of data sets in either hardcopy or softcopy);

2) Direct to data store service (e.g. software delivery via FTP, specified via e-commerce order request);

3) Brokered service: provide specification of data access request to secondary (online or offline) access service;

4) Online data service (e.g. request/response access protocol to data warehouse) supporting online operations such as:

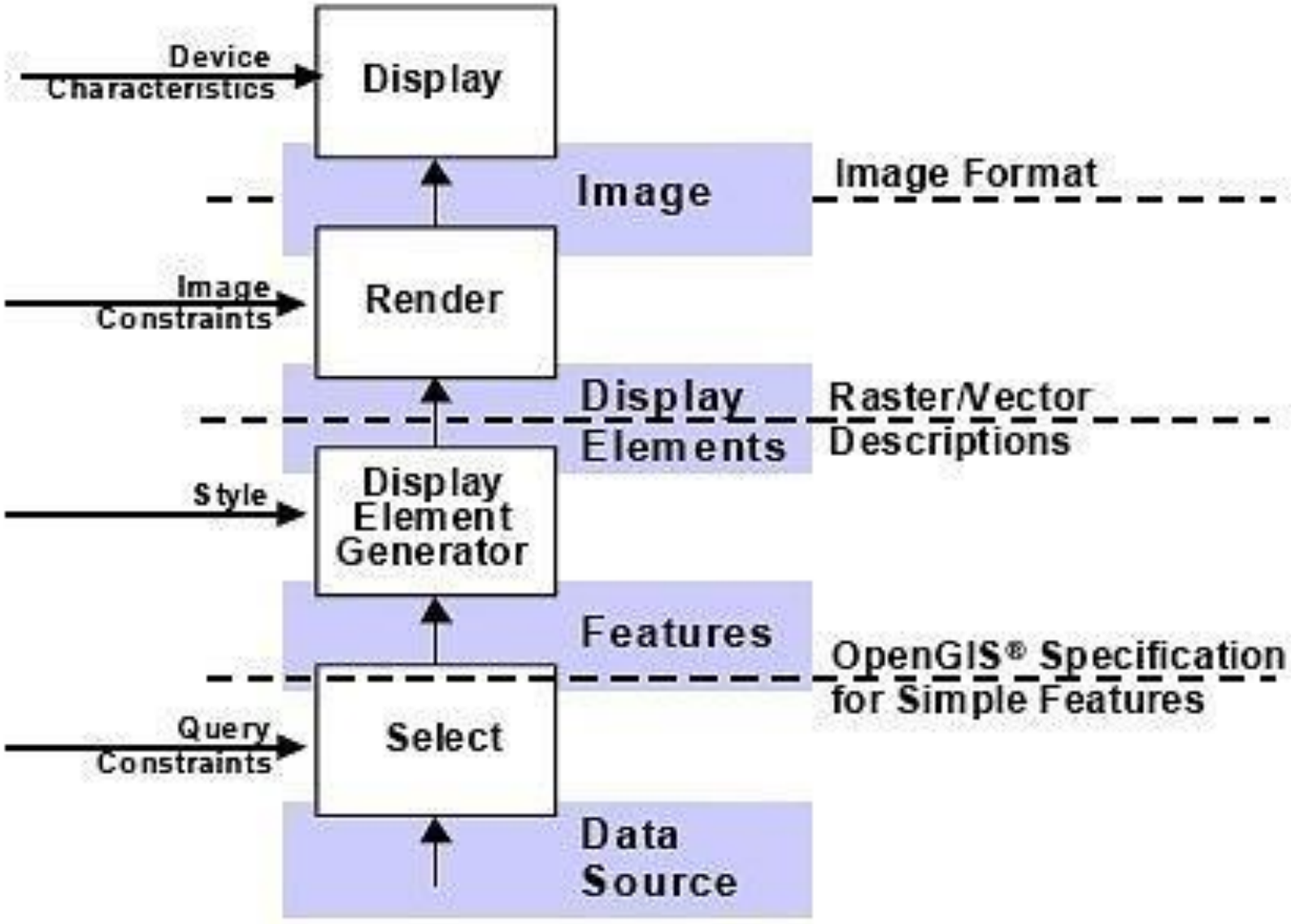
a) Drill down;

b) Aggregation;

c) Generalization.

2. In Open Geospatial Consortium Project Document 98-060: "User Interaction with Geospatial Data" **the OGC Portrayal model** is described. *Figure 6.1* describes this model, which illustrates a simple features-based access and portrayal (*presentation of information to humans*) services pipeline.

Fig.6.1 – OGC Portrayal model



6.4.5. SDI data access clients

Online implementations of data access clients include:

1) **'Thin client'**, which is provided by standard Internet/Web tools (no Java – e.g. Web browser, e-mail, FTP client, etc.);

2) **'Medium client'**, which is provided by Web browser with Java, or ActiveX controls;

3) **'Thick client'**, which is provided by a Web browser plugin, or standalone application (network access via a distribution computing platform such as CORBA, Java RMI, etc.).

6.4.5. SDI data access clients:

4) **Traditional GIS type client**, which need access to previously downloaded data set, and direct network access to data warehouse;

5) **'Middleware client'**, which need transparent access to consumer via a middleware infrastructure or applications service;

6) **Geoprocessing service client** – direct access to data for use by a geoprocessing service (e.g. Web Mapping with interactive portrayal service).

6.4.6. SDI data formats

Common spatial data formats used in SDI include the following.

1. **GIS proprietary formats** (e.g. ESRI, MapInfo, Intergraph, etc.):

– A good overview of GIS formats can be found at *http://www.gisdatadepot.com/helpdesk/formats.html*

2. **International and community formats**:

1) Efforts have recently been made to minimize the number of geodata formats and to converge towards a reduced set;

2) Examples of this are:

a) The Spatial Data Transfer System (SDTS),

3) There are also exchange formats that allow the use of data outside of closed environments:

- E.g. Geography Markup Language.

4) Due to lack of consensus on specific format standards, spatial data infrastructures often support access to multiple spatial data formats through data access services;

5) However, if it is feasible, the definition of a single community format based on ISO and OGC specifications is ideal to promote information exchange (See Theme 2);

6) Currently, most GIS and related access systems support format translation. Examples of commercial support for format translation include:

a) The Feature Manipulation Engine from Safe Software (*http://www.safe.com/*);

b) Geogateway from PCI (*http://www.pci.com/*);

7) An online data access service that combines data access with format translation is the Open Geospatial Datastore Interface (*http://ogdi.sourceforge.net*).

3. **Web implementation formats**

Vector files:

1) There are a *three candidate file formats* for encoding vector information on the WWW:

a) Simple Vector Format (SVF);

b) Web Computer Graphics Metafile;

c) XML-based encoding formats, e.g. GML;

2) Only GML is specifically designed for the encoding of vector geographic information;

3) The other formats are designed for the exchange of vector graphic information but may have little or no reference to real world or mapped coordinate systems or feature

Raster files:

1) Web/internet delivery of GIS raster formats such as ADRG, BIL and DEM is often problematic due to the large size of such files, combined with general lack of Internet bandwidth;

2) Typically compressed raster files predominate Web-based portrayals;

3) Common compressed Web formats include GIF, JPEG and PNG to move single variable panchromatic or color images as raster files.

6.5. SDI data access's relationship to other SDI services

1. ***Figure 6.2*** illustrates the relationship role of data access in an end-to-end **spatial resource discovery, evaluation and access paradigm**.

2. Successive iterations of resource discovery via a metadata catalogue, followed by resource evaluation (such as Web Mapping) lead to data access either:

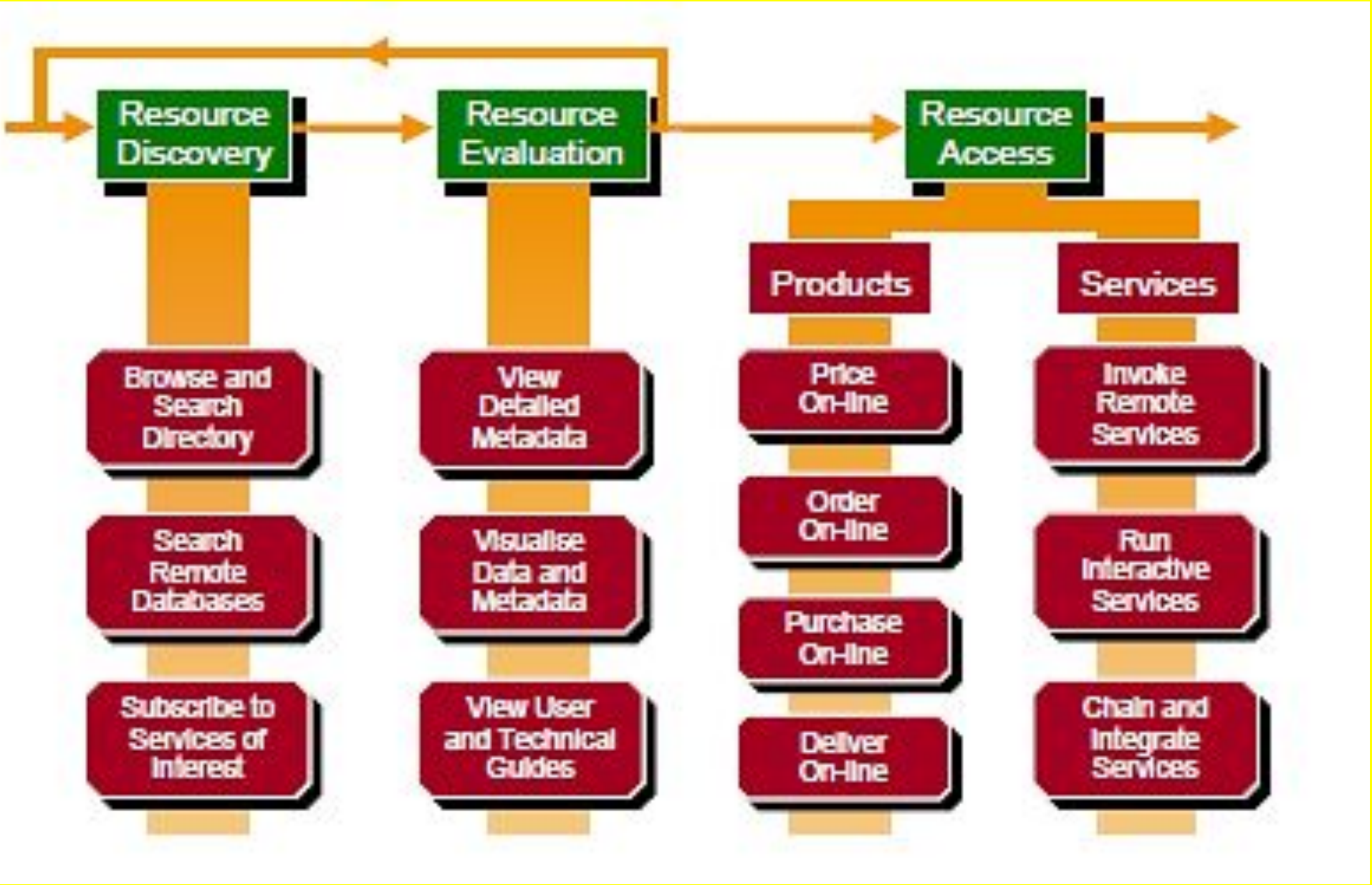
   1) Direct as a data set;

   2) Indirect via a data access service.

3. Mature spatial data infrastructure will allow both application and human exploitation of the resource access paradigm.

4. A key element of future spatial data infrastructures is the

Fig.6.2 – Spatial resource discovery, evaluation and access paradigm



5. Future capability for chaining of distributed geoprocessing services is also expected.

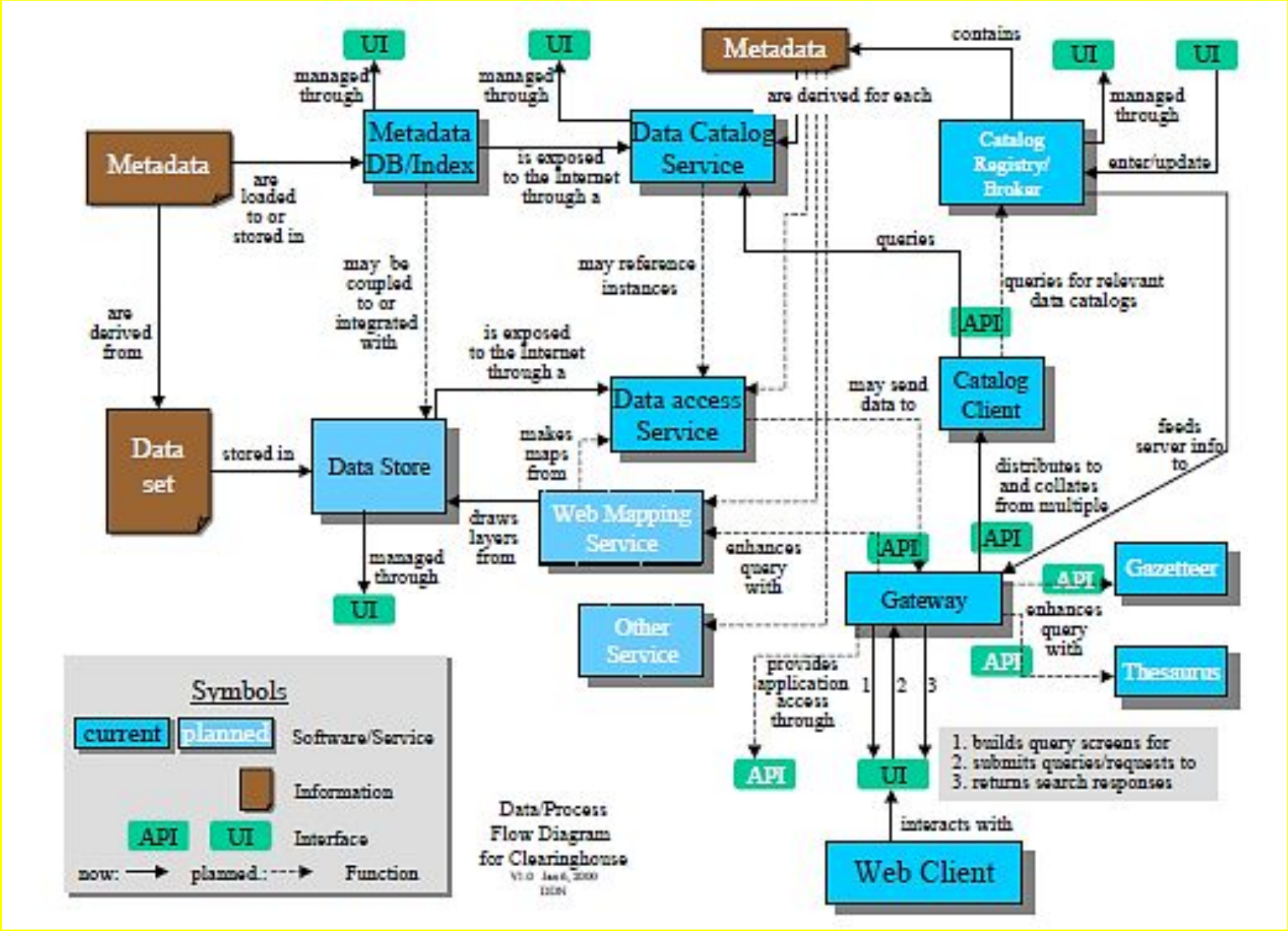
6. A **system context for spatial data access services** is given in ***Figure 6.3***:

1) A data access service provides network access to a data set stored within a data store;

2) Data sets are discovered (and later accessed) via metadata queries from a catalogue client to a data catalogue service;

3) Data sets can be visualized (and later accessed) via Web Mapping Services, which are complementary to the Data Catalogue Service.

Fig.6.3 – System context for spatial data access services



6.6. SDI data access standards

In general, standards related to spatial data access are still in their infancy.

The standards of most relevance to access components of spatial data infrastructures include those from:

1) ISO/TC211;

2) Open GIS Consortium (OGC);

3) Internet-related bodies including:

a) The World Wide Web consortium (W3C);

b) The Internet Engineering Task Force (IETF)

6.6.1. ISO/TC211 data access standards

1. The primary mandate of ISO/TC211 is international standardization in the field of digital geographic information.

2. These ISO standards may specify, for geographic information, methods, tools and services for data management (including definition and description), acquiring, processing, analyzing, accessing, presenting and transferring such data in digital/electronic form between different users, systems and locations.

3. Work on such services is currently underway in both ISO/TC211 and the OGC.

4. The definition of services interfaces will allow a wide range of applications access and use of spatial resources.

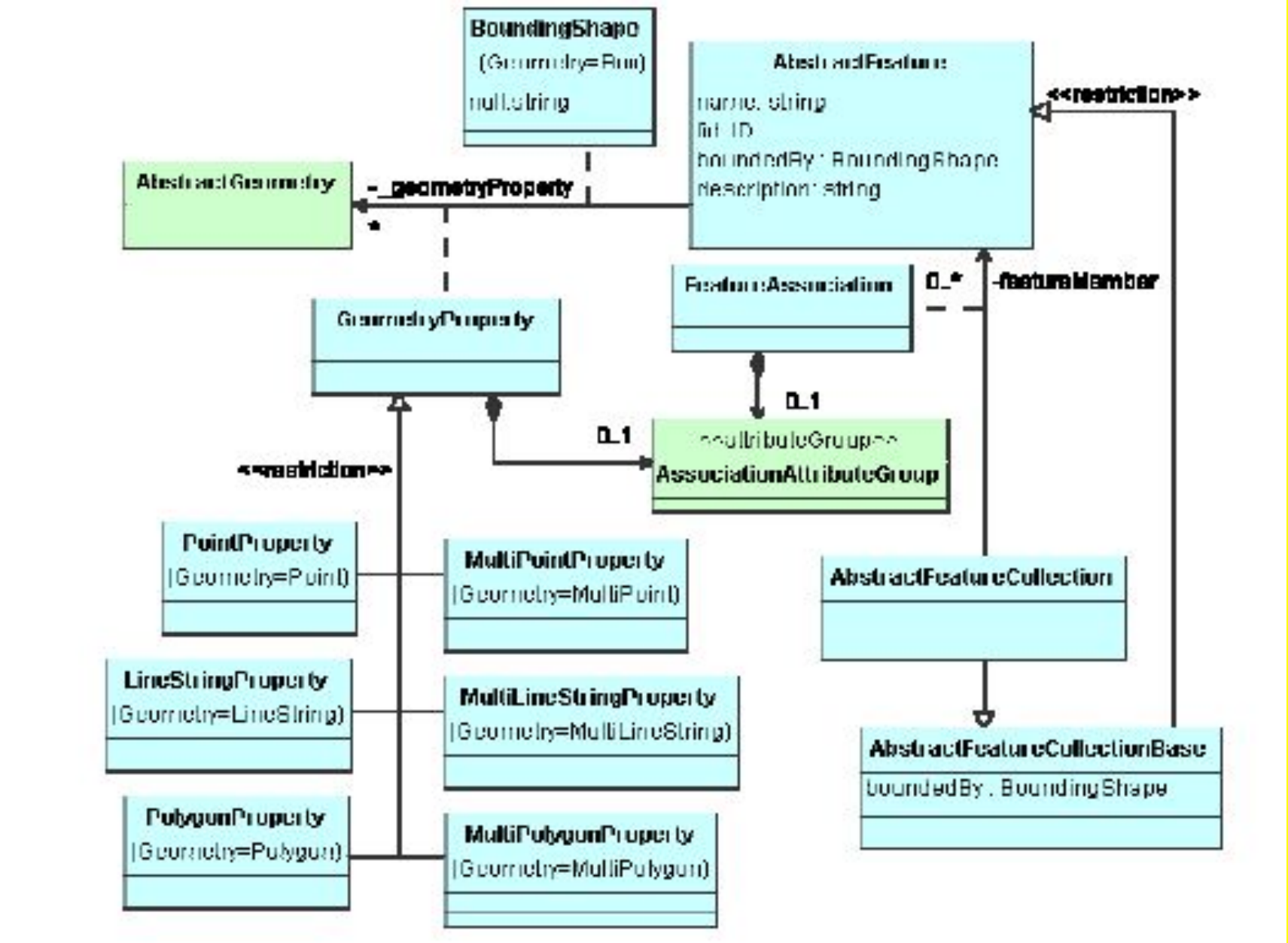
6.6.2. Open GIS Consortium (OGC) data access specification

1. The Open GIS Consortium has achieved consensus on several families of interfaces, and some of these have now been implemented in Off-The-Shelf software.

2. The publication of the **OGC Web Feature Service (WFS) Specification** in 2002 provided a solution for the standardized request and delivery of **vector data**.

3. Supporting the **OGC Feature Model** shown in *Figure 6.4*, the WFS specification defines the dialogue required to interact with geographic features via vector data service.

Fig.6.4 – UML Model of the OGC Feature Model



4. **Web Feature Service (WFS)**:

1) A service that can describe data manipulation operations on OGC simple features (feature instances) such that servers and clients can "communicate" at the feature level;

Note. **Simple feature** – feature restricted to 2D geometry with linear interpolation between vertices, having both spatial and non-spatial attributes.

2) A Web Feature Server request consists of a description of the query and data transformation operations that are to be applied to WFS Web-enabled spatial data;

3) The request is generated on a client and is posted to the WFS Server.

4) The WFS Server interprets the request, checks it for

5. **The OGC Web Coverage Service (WCS) Specification** was published in 2003.

6. It extends the Web Map Service (WMS) interface to allow access to spatial "coverages", rather than WMS generated maps (pictures).

7. **Web Coverage Service (WCS)**:

1) A service that supports the networked interchange of spatial data as coverages containing values or properties of geographic locations;

2) The WCS provides access to intact (unrendered) spatial information, as needed for client-side rendering, multi-valued coverages and input into scientific models and other clients

3) **Coverage** is a feature that acts as a function to return values from its range for any direct position within its spatiotemporal domain:

a) A coverage represents **continues phenomena**;

b) There are different types of coverage – a set of tiled polygons, a grid of values, a mathematical function, or a combination of these;

c) Grid coverage represents the value in a grid's points;

4) WCS provides receiving an array or surface of data values.

8. **The OGC Web Coverage Service (WCS) Specification** have also been released to support feature access in relational database environments: one each for SQL, COM-based, and CORBA distributed computing platforms.

9. The WCS interfaces provide access to and control over GIS features **at 3 levels**:

1) At **the primitive level**, the interfaces provide for the establishment of linear and angular units, spheroids, datums, prime meridians, and map projections that give semantics to coordinates;

2) At **the intermediate level**, they enable the construction and manipulation of geometric elements such as points, lines, curves, strings, rings, polygons, and surfaces, as well as the topological and geometric and other relationships between them;

6.6.3. Web and Internet related data access specifications

1. **The World Wide Web consortium** or **W3C** is responsible for the development of common protocols and specifications to further the evolution of the World Wide Web.

2. Activities of the W3C that related to spatial data access include work on Web graphic file formats, XML and metadata.

3. **The Internet Engineering Task Force (IETF)** develops and maintains specification for many Internet related application, transport, routing and security standards (RFCs) many of which are related to data access.

6.7. Best practice application of SDI open data access

1. One common problem with online access to data is the variety of policies and practice in place by the different data custodians, which use **different basic paradigms**:

1) Custodians who restrict access to particular users would benefit from common user authentication/authorization services;

2) Custodians who charge for data or services would benefit from electronic commerce services;

3) Custodians who distribute data free of charge would benefit from an inexpensive mechanism (both time and

2. One example of services to support the third paradigm is **GeoGratis** (*http://geogratis.cgdi.gc.ca/*) that provides common services to support the distribution of freely available geospatial data.

3. GeoGratis provides a single FTP/web access point where consumers can discover and download freely available data sets.

6.8. Evolution of data access and related spatial data services towards SDI

1. The matrix below (*Fig.6.5*) illustrates the evolution of data access and related spatial data services.

2. Migration from “classic” towards “infrastructure enabled; standards based; and full functioned” is required to bootstrap a national spatial data infrastructure.

3. Early adoption and “best practices” of such evolution should be followed by key government data providers.

Fig. 6.5 – Evolution of data access and related spatial data services towards SDI

	Classic	Move to online	Infrastructure enabled; Standards-based; Full functioned
Metadata	Ad hoc	FGDC - based	ISO TC211 - based
Metadata Catalogue	Offline, hardcopy Compact disk	Database enabled; Web accessible	Semantic interoperability via search/retrieval protocol OGC catalogue
Visualisation	Offline: fax, hardcopy, Compact disk	Web – accessible, Map enabled	Visual evaluation via OGC WMS
Ordering	Phone, fax	E-mail	Web-based, integrated with e-commerce payment
Product selection	Predefined products	Geographic and layer-based subsetting of predefined products	Selection of arbitrary features, layers and feature collections from seamless data warehouse, using OGC WFS and Filter Encoding
Delivery	Offline: hardcopy	Offline: softgood (e.g. Compact disk)	Online: File-based for network download (note: file may be generated dynamically) OGC WFS
Packaging/ formatting	Offline: hardcopy or softgoods from predefined formats	Online: user specified format selected from pre-generated softgoods	Online: support for user-specified softgood format via dynamic format translation OGC GML
Payment	Offline: traditional consumer	Online credit-based payment to registered list of consumers	Online e-commerce based, supporting "previously unknown" customers (e.g. online credit-card payment)