Introduction to Spatial Data Infrastructure (SDI)

SDI?

The explosive growth of the Internet continues to revolutionize the way modern day business is conducted and services provided. In recent years geographical information systems (GIS) and enterprises have continued to evolve towards distributed models in order to better exploit the potentials presented by the Internet computing paradigm. GIS systems have exhibited sustained evolution from stand alone, data-centric stovepipes to distributed models composed of open interoperable services while GI enterprises continue to pursue edible models in order to leverage advances in business networking and e-commerce. Meanwhile the spatial data infrastructure (SDI) concept which emerged in the 1980s to advance spatial data sharing by taking advantage of the ubiquity of the Internet and its ease of use has matured and is evolving into an infrastructure for the delivery of geoprocessing services, the so called geographical services infrastructure (GSI).

The term "Spatial Data Infrastructure" (SDI) is often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data. The SDI provides a basis for spatial data discovery, evaluation, and application for users and providers within all levels of government, the commercial sector, the non-profit sector, academia and by citizens in general.

Spatial data infrastructure encompasses the resources, systems, network linkages, standards and institutional issues involved from many different sources in delivering geo-spatially related information to the widest possible group of potential users. Also, adopting this definition with general definition of infrastructures, it could be concluded that SDI has following characteristics:

- Is a set of base capabilities;
- Is a general comprehensive system;
- Has multiple effective aspects in ICT utilization;
- Includes some stable and dynamic physical components;
- Provides important, fundamental and irreplaceable services.

SDI Evolution

1906: Created U. S. Geographic Board.

1919: Created a new Board of Surveys and Maps that took over the responsibilities of U.S. Geographic Board.

1953: Describes responsibilities of Federal agencies with regard to the coordination of surveying and mapping activities.

1967: Better describes responsibilities of Federal Agencies (leadership and coordination).

1983: Establishes coordination of Federal digital cartographic data programs.

1990: Establishes Federal Geographic Data Committee and links more programs.

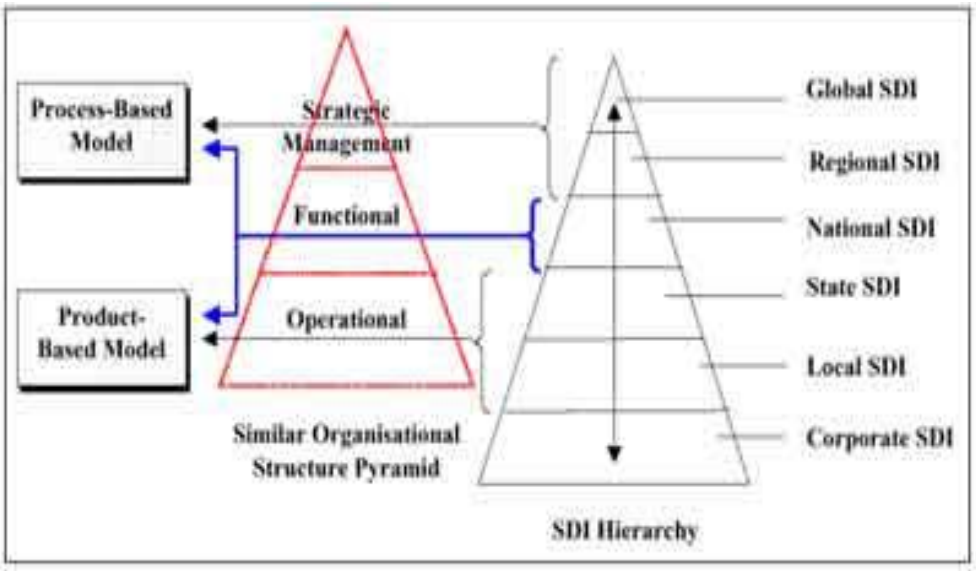
1994: Establishes the National Spatial Data Infrastructure (NSDI).

An infrastructure is a kind of organization, which is the main basis for other organizing activity developments, contributes necessary and different activities in sustainable development. In this process new and adoptive nature is created for each activity, preserving their usual specification through strategies and policies injection and integration assuring their effectiveness.

Tracing SDI trend, after one decade of its emergence (1990s), its concepts have been fulfilled considerably. Balancing the GI (Geographic Information) generation and utilization concentration is the main evolution resulted new SDI definitions as follow:

- Product-based model, which represents one of the main aims of an SDI development initiative, can be used to link existing and upcoming databases of the respective political/administrative levels of the community.
- Process-based model, which presents one of the other main aims of an SDI development initiative, defining a framework to facilitate the management of information assets.

Also, a hierarchical structure is defined for SDI, which comes from especial nature of geo-spatial data as they depend on scale and application. Regarding this structure, SDI treats in a pyramid, which its base (corporate level) is very similar to current foundation used for desktop GIS applications development and by its promotion to the peak, SDI converts to a global infrastructure (Figure below).



Considering this structure, the following points could be illustrated:

- From each level we could navigate to its upper and lower levels.
- Each lower level provides the building blocks of upper levels.
- Each upper level transfers its overall definitions and backgrounds (strategy, policy) to lower level;
- The levels relationship fastens as they come closer.

These specifications and tangibility of this structure for societies, has clarified and simplified SDI concepts and developments more.

What is a National Spatial Data Infrastructure (NSDI) ……?

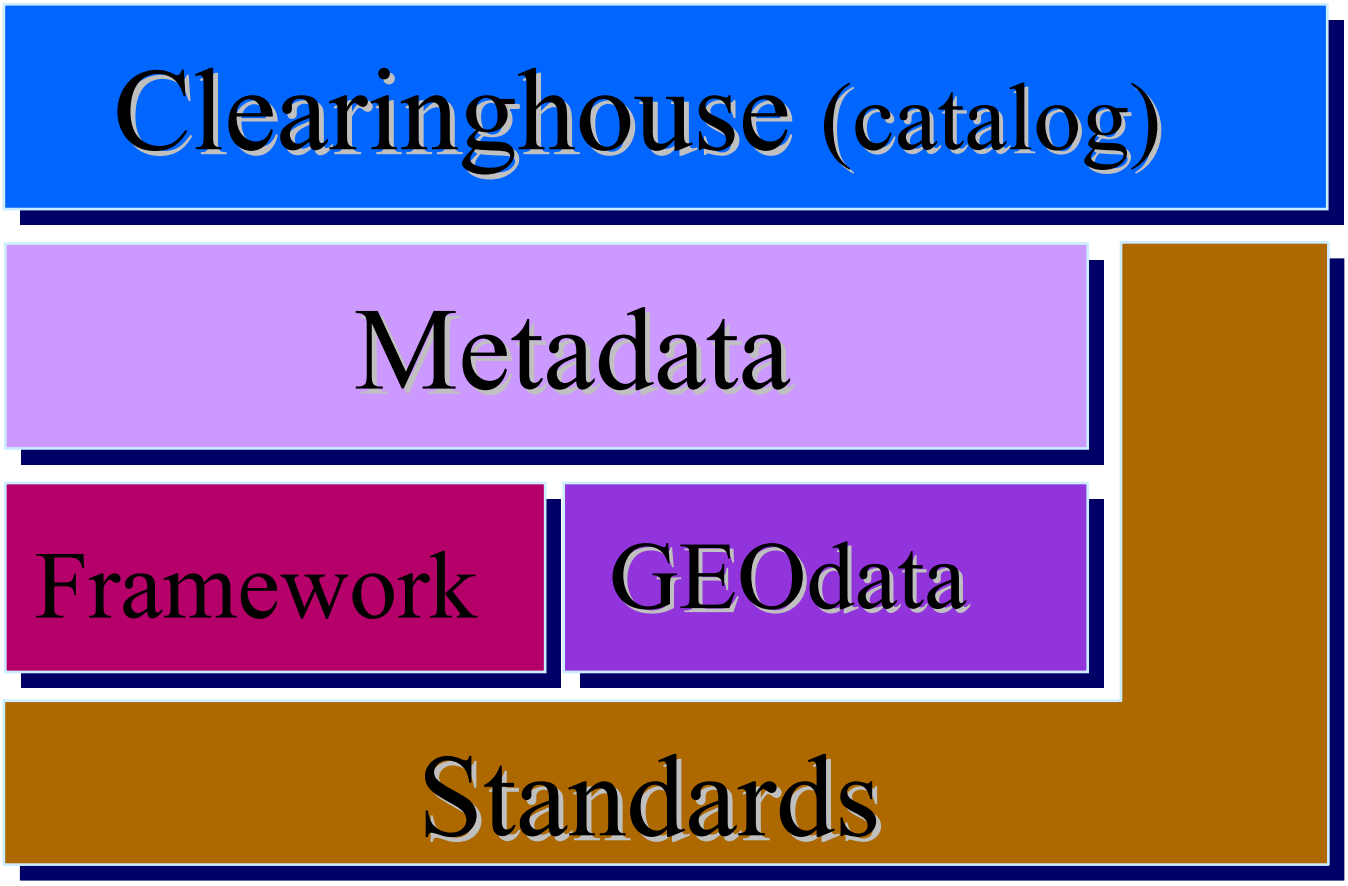
To encourage the collection, processing, archiving, integration, and sharing of geospatial data and information using common standards and interoperable systems and techniques

….. and accessible via the web.

<u>The Vision of the NSDI:</u>

To assure that spatial data from multiple sources (Federal, State, and local governments, academia, and the private sector) are widely available and easily integrated to enhance knowledge and understanding of our physical and cultural world.

Components of SDI



Clearinghouse

A (spatial data) clearinghouse is a distributed network of spatial data producers, managers and users that are linked electronically together. It is a system of software and institutions that are to facilitate the discovery, evolution and downloading of digital spatial data and provides means to inventory, document and data sharing. The clearinghouse concept is a useful one in building a Geographic Information Infrastructure (GII). The objective is to minimize unnecessary duplication of effort for data capture, and to maximize the benefit of geographic information sharing.

Data providers nowadays are fully aware of the importance of advertising and making available their metadata describing their databases, to facilitate the use of their products. This explains the current level of activity of building these clearinghouses.

Data Clearinghouse!!

- Institutional view
  People and infrastructure to facilitate discovery of who has what geographic information.
- Technical view
  A set of information services that use hardware, software, and telecommunications networks provide searchable access to information.

How does a clearinghouse work?

A clearinghouse allows data providers to register their geographic data sets, the quality of these data and also the instructions for accessing them. Each data provider provides an electronic description of each spatial data set. In addition, the provider may also provide access to the spatial data set itself. The clearinghouse thus functions as a detailed catalogue service with support for links to spatial data and browsing capabilities. The data described in the clearinghouse may be located at the site of the data producers or at sites of designated data disseminators located elsewhere in the country. Obviously computer network facilitates are the key factors to success.

Metadata Concepts and functionality

Metadata is defined as background information that describes the content, quality, condition and other appropriate characteristics of the data. So metadata is a simple mechanism to inform others of the existence of the data sets, their purpose and scope. In essence metadata answer who, what, when, where, why and how questions (WH Questions) about all facets of the data made available.

Metadata can be used internally be the data provider to monitor the status of data sets, and externally to advertise to potential users through a national clearinghouse. Metadata are important in the production of a digital spatial data clearinghouse, where potential users can search for the data they need.

Roles of metadata:

Applicability: information needed to determine the data sets that exists for a geographic location,

Fitness for use: information needed to determine whether a data set meets a specified need,

Access: information needed to acquire an identified data set,

Transfer: information needed to process and use a data set,

Administration: information needed to document the status of existing data (data model, quality, completeness, temporal validity etc…) to define internal policy for update operations from different data sources.

Metadata Standards

For metadata to be easily read and understood, standards create a common language for users and producers. Metadata standards provide appropriate and adequate information for the design of metadata. The key development in metadata standards are the ISO STANDARD 1504615 METADATA, the federal geographic data committee's content standard for Digital Geospatial Metadata FGDC.

A standard provides a common terminology and definitions for the documentation of spatial data.

Metadata management & update

Just like ordinary data, metadata has to be kept up-to-date. The main concerns in metadata management include what to represent, how to represent, how to capture and how to use it; and all these depend on the purpose of metadata;

For internal (data provider) use, we will refer to local metadata which contains the detailed information about data sets stored on local hardware and managed by the data provider. For external use we refer to global metadata which contains a short description of the data sets (an abstraction of local metadata) as advertised in the clearinghouse to allow users to find relevant data efficiently.

Data providers should register their data holding with the clearinghouse. Whenever changes occur in their data, each data provider reports the changes to the clearing authority. Updating the global metadata is the responsibility of clearinghouse.

Data sharing and related problems

Geographic data exchange and sharing means the flow of digital data from one information system to the other. Advances in technology, data handling and data communication allow the user to think of the possibility of finding and accessing data that has been collected by different data providers. Their objective is to minimize the duplication of effort in spatial data collection and processing. Data sharing problems which can be viewed as critical factors in SDI can be briefly described as follows:

1. *Data standards*: It refers to an agreed upon way of representing data in a system in terms of content, type and format. Exchange of data between databases is difficult it they support different data standards or different query language. The development of common data

architecture and the support for a single data exchange format, commonly known as *standard for data exchange* may provide a sound basis for data sharing. Examples of these standards are the DIGEST (Digital Geographic Information Exchange Standard) Spatial Data Transfer Standard (SDTS) etc...

2. *Heterogeneity:* it means being different in kind, quality and character. Spatial data may exists in a variety of locations, are possibly managed by a variety of database systems, were collected for different purposes and by different methods, and are stored in different structures. This brings about all kinds of inconsistency among these data sets (heterogeneity) and creates many problems when data is shared.
3. *Communication problems*: with advances in computer network communication and related technology, locating relevant information in a network of distributed information sources has become more important recently. The question is which communication technology is the best suitable for transfer of huge amounts of spatial data in a secure and reliable way. Efficient tools and communication protocols are necessary to provide search browse and delivery mechanisms.
4. *Institutional and Economic problems*: these problems arise in the absence of policy concerning pricing, copyright, privacy, liability, conformity with standards, data quality etc… resolving these problems is essential to create the right environment for data sharing.