

## Spatial Data Infrastructures (MS No. 39)

### Glossary

**Data Protection:** The requirement that data about people conforms to rules of non-disclosure and privacy.

**GeoSpatial information** is any information that can be geographically referenced, i.e. describing a location or any information that can be linked to a location.

**Global Positioning System (GPS).** A satellite-based navigational system allowing the determination of a unique point on the earth's surface with a high degree of accuracy given a suitable GPS receiver.

**Interoperability** is the ability to link different applications regardless of proprietary software used at the source together so that several web services are available concurrently.

**Metadata:** the data describing the data. Typically this consists of a number of descriptor fields identifying content, quality, currency, ownership and other characteristics.

**Standards:** an agreed set of minimum requirements, which are to be reached by all data providers.

### Introduction & Context

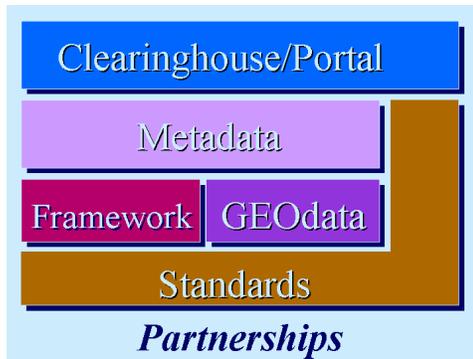
With the development of Geographic Information Systems from the late 1960's on and with the rapid development of computing technology and power, geographical information (GI) and digital spatial data have become key resources in all branches of geography. As cartographic data have become increasingly digital, both in terms of their sourcing and manipulation, the power of spatial data is increasingly recognised across the globe. The recent arrival of cartographic resources into mainstream consciousness through the development of Google Earth and other Internet based data has also created an increased awareness of the power of mapping. At a governmental level the increasing use of GI in all branches of government and its widespread adoption and use in civil society is well established in the

developed world and is growing in the developing world as well. Within society as a whole there is also an increased awareness of the ways in which spatial data underpin society in the form of in-car navigation systems, satellite imagery use in the media and the use of GI in business demographics. One of the principal tasks of GI experts therefore, is to manage and shape this increased potential and awareness and one of the key global models used to try and do this are Spatial Data Infrastructures (SDI).

One definition of a Spatial Data Infrastructure (SDI) is:

*The technology, policies, standards, and human resources necessary to acquire, process, store, distribute, and improve utilization of geospatial data.* (Source: FGDC 2004)

SDI's were first discussed and proposed in the early 1990s in both Europe and the US. The influential report by Lord Chorley published in London in 1987, focused on the potential of spatial data to underpin much of the future informational needs of society. Specifically, Chorley made a number of recommendations around the human dimension of GI and the need for a co-ordinating body to manage the development of the industry and the data. This recognition of the need for co-ordination was particularly pertinent to discussions on copyright, standards, interoperability and take-up. The co-ordination role to fully develop and explore the potential of GI is one that lies at the heart of SDI. In the early 1990s a similar initiative in the US, driven by the Federal Geographic Data Committee (FGDC), led to the formal endorsement by the then President, Bill Clinton to develop a national spatial data infrastructure (NSDI). Figure 1 below outlines the core components envisaged in the US NSDI.



(Source: FGDC <http://www.fgdc.gov/components>)

One of the leading proponents of SDI is Ian Masser, who saw spatial data as a national asset, having something in common with other infrastructural elements such as railways and roads. The spatial scale of operation for SDI is important and much attention in recent years has been focused on the need for regional, national and even global organisations all of which are linked by common standards but which also vary from continent to continent and country to country. Masser also identified the national SDI (NSDI) as the most developed level with other levels identifiable at super-national (GSDI and Inter-Continental versions) and sub-national level (typically in federal countries). At the beginning of 2007, there were over 120 countries worldwide either actively developing, or considering the development of, an SDI.

### Key Components of an SDI

There are a number of key components that underpin SDI, whatever the spatial scale at which it operates. These include legislative frameworks, content and data frameworks, metadata and standards, quality issues, geo-portals, human factors and dissemination and implementation. By definition, each of these are diffuse and complex areas but to realise a vision of an SDI, they are all essential in teasing out the real life issues that such an initiative faces. They also need to be understood in placing SDI more closely within the wider GIS and geography disciplines.

*Legislative Frameworks:* The development of any SDI requires a minimum level of legal standing and support. The relationship between GIS and the Law is still being worked out and varies considerably across the globe. This depends also on the geographical scale of the framework and can range from the local and regional to the national, international and global. Initiatives at world (Global SDI) and continental level (e.g. the EU INSPIRE initiative) aim to set these agreed legal frameworks in place. At a spatial data level, agreements on data protection, copyright and ownership all have the potential to be serious barriers to the development and adoption of SDI and need to be addressed at a very early stage. It is clearly easier to handle national legal requirements than international ones

*Content and Data Frameworks:* The range and coverage of spatial data are extensive at the beginnings of the 21<sup>st</sup> century. Many subjects collect and use spatial data and these range from administrative data routinely collected by government on employment and electoral rights to subject-specific data such as health care utilisation and land ownership. The level of coverage does however vary considerably both between and even within countries. While SDI do not aim to suggest an agreed standard minimum data set, they provide instead a framework or template into which sets of digital spatial data may be arranged and organised to an agreed international standard. Agreements on what core data layers should be included in these frameworks is a source of considerable debate within SDI. The US NDSI lists seven framework themes, which include: geodetic control, cadastral, orthoimagery, elevation, hydrography, administrative units and transportation. These provide a good working model for most SDI.

*Metadata and Standards:* As part of the development of spatial data standards, in part driven by the Chorley Report, a series of technical committees, namely 211 and 278, have met at ISO level to develop and agree a set of technical standards and

agreed metadata formats for spatial data. Current standards include the ISO19100 Series/TC 211 as well as standards developed by agencies such as the FGDC in the US and the private sector Open GeoSpatial Consortium (OGC). Metadata which, simply put, is 'data about data', is essential to these standards in that it requires data holders to use standard descriptors to document the precise nature of the spatial data holdings. A number of international standards such as the Dublin Core are already being used in this area.

*Quality Issues:* Clearly, it is impossible to manage and standardise spatial data from all over the globe. Quite apart from difficulties with projections and coordinates, national and private mapping agencies produce maps at a bewildering range of scales, few of which are standard. Similarly, the derivation and accuracy of digital spatial data varies hugely around the world. Even in countries with a long scientific cartographic tradition such as the UK, Germany and the US, the arrival of new technologies such as satellites and GPS have brought with them full adjustments to the national grids. Much of the initial work of global SDI has focused on the development of initial small-scale but worldwide layers such as the Digital Chart of the World, a global dataset released in 1993. Other quality issues relate to scale, currency and the breadth of attribute data.

*Geo-Portals & Distributed Information Systems Architecture:* A key aspect of any SDI is that, once data has been collected, collated and verified to agreed international standards, it must be disseminated to potential users. Quite apart from the complex legal aspects of data dissemination and data sharing, the technology must be capable of allowing spatial data to be distributed equally, correctly and with an acceptable level of ease for the end user. The growth of the Internet has been a huge factor in making such Geo-portals feasible for the distribution and on-line cataloguing of spatial data sets. Geo-portals are a generic name for the IT service locations which users access to identify, view and

download spatial data sets. As such they are the 'public face' front-end to any SDI and are a key outcome of developments in 'Internet GIS'.

*Human Factors:* As the quantity and quality of spatial data has mushroomed since the 1980's, the value and importance of spatial data has also been recognised. In an increasingly IT literate world, information and access to it brings with it associations with power and financial gain. As such, the ownership of data, who is allowed use it and who should benefit from it are key concerns. Linked to this, the rights of individuals to privacy and the wider importance of data protection offer a real challenge to SDI's worldwide. As the fundamental aim of SDI is to pull together and disseminate spatial data sets, the potential for individuals and agencies to prevent this happening is enormous for a variety of reasons including, national interests, legal challenge, work priorities and commercial concerns.

*Dissemination and Implementation:* Quite apart from the effective dissemination mentioned above, the real challenge to full SDI development lies in its implementation. Much of that implementation is linked to both data audit (on the collection side) and public access (on the dissemination side). A good SDI should by definition have a minimum core set of data, at an acceptable scale and with an acceptably complete level of coverage. All of these elements are easier said than done.

## **Current Spatial Data Issues**

This section will outline current research around a range of key spatial data infrastructural issues, including standards for both data and metadata, interoperability versus harmonisation, data quality, access, intellectual property rights (IPR) and digital rights management, and managing change. All of these issues have an impact on how spatial data is collected, processed, stored, interpreted, used, and disseminated by the owning organisation and how it is accounted for by data custodians (Groot and McLaughlin, 2000).

Much of the research on SDI has been driven by technical discussions around standards, inter-operability and dissemination. The work of Burrough and Masser in Europe and by the FGDC in the US had been mirrored by work in Africa, Asia and Australasia. Much of this work has yet to filter into mainstream geographical literature though the work of Pickles in creating a debate on the human implications of GIS may yet be replicated as the wider world picks up on the potential of GI and its organisation within SDI's. In particular critical debates on data ownership, power and representation are also appearing in the wider human geography literature but these may be necessary to fully understand the implications of SDI development and to help clarify where ISDI can really work for the common good in terms of opening up access to spatial data and promoting it's optimal utilization.

The literature on Standards and Interoperability is best summarised in some of the current literature. Masser, in particular, identifies the current state of play around the world in terms of national, regional and global standards and provides a selective history of their development. He comments in particular on the recent organisation of a global initiative, GSDI that is a natural progression from the smaller national level SDIs. Issues of data quality continue to be crucial to the acceptability and success of SDI development. Evidence suggests that public expectations have increased with the IT revolution and such expectations are clearly linked to acceptable standards in terms of the level of accuracy, the number of available datasets and usability of spatial data sets collected and disseminated within SDI frameworks.

Data access is a subject that has always exercised geographers and despite the obvious efficiency gains in storage and maintenance in a digital age, there are still trade-offs taking place between data owners and data users. The dissemination of spatial data within SDI is in part driven by an understanding of public rights to

spatial data. This clearly varies internationally in that different states have different attitudes to the release of spatial data, ranging from free public access to heavily protected and costly data sets. Some of these difficulties also reside in debates around IPR and 'added value'. Arguments abound as to the fact that raw spatial data is developed by states but the value-added product is the comprehensible spatial map layer which the user needs and the ownership of that is vested in the data creators/owners. Increasingly, privately collected spatial data, enhanced by technologies like GPS and Lidar, are an additional complication in relation to ownership and access.

### **Current Drivers and Barriers**

While much of the literature and discussions around SDI have been embedded in technical discussions and the wider policy and political implications, it is worthwhile to summarise the current state-of-play in terms of Drivers and Barriers. An implicit aim of SDI is to place in the public arena a set of spatial data layers which can be easily accessed, used and manipulated. It is almost certain that the Internet will be the medium for this dissemination and that this in turn may allow for potentially free and unrestricted use of those datasets. This outcome is what should ultimately guide how SDI is developed. If there are concerns and difficulties these will determine the nature of the data finally provided and the one imponderable at this stage is what form the access and utilization of that data will take. Will access to such data create a demand for more? Will the ways in which the public interpret and use that data affect what data is provided? Will the data be exploited commercially by the data users more than the data providers? Will SDI underpin global developments in e-government in a positive or negative way?

Arguably, once users understand and use distributed spatial data sets, as envisaged by SDI, the flexibility and sheer ease of access should arguably create an additional demand. The success of

GeoPortals such as the Alexandria Digital Library and the National Statistics website in the UK show just how such access has enabled individuals, community groups, agencies and other government departments. With the development of the internet and more significantly the growing use of spatial data in normal daily life (SatNav, GPS, Weather maps, Google Earth), there is evidence for a growing voluntary spatial data creation process which currently lies parallel to any putative SDI frameworks. For many data providers, a continuing concern is the way in which their data might be misused or misinterpreted.

Spatial data has enormous potential in maximising economic, social and environmental benefits. The commercial, social and economic value of such datasets are being recognised daily. A conflict may exist between the raw data and its use for commercial profit or state control. One thing that should be pointed out is that SDI make no explicit provision for free access to spatial data, though some people may understand it as this. Essentially, there is the scope in SDI to disseminate data in a charging regime that may range from free to full commercial cost. Getting the pricing right is one of the challenges which may affect the success of SDI development.

At a wider level, SDI has the potential to provide a good example of technology transfer, bringing the technical knowledge of the spatial data community into the real world for its' greater benefit, while at the same time safeguarding the standards of that data. Other potential benefits relate to the development of agreed models and templates to improve data sharing protocols and also provide a form of future proofing as volumes of available data increase. The global initiatives around e-government are also very strong drivers for the implementation of SDI as they provide a technical and political solution in disseminating government data.

Barriers to the successful implementation of SDI are manifold. Some of the difficulties encountered in developing the

US NSDI are well known. Interestingly standards and inter-operability aspects have proved much less difficult than agreements on what datasets should be included and the difficulties of working with multiple stakeholders. This latter aspect, which incorporates legal and human factors associated with data sharing, data access and both national international agreements continues to hold up progress in a number of key area. Additionally, the continuing speed and unpredictability of ICT technological change through the development of location-based services and mobile technologies represent both a barrier and an opportunity. Data access and data creation are enabled through such technologies. Linked to this the awareness of spatial data through societal advances in GPS/Mobile applications may embed the power and value of spatial data even further into human consciousness. Such developments also have the potential to spiral out of control and perhaps a full and concretized development of SDI may be one way to manage this potential for the common good.

### **The Future of SDI**

Finally, there are a number of ongoing discussions and large-scale research initiatives about the ways in which SDI may develop. The Australian SDI (ASDI) identifies five priority areas, namely; Governance, Access to Data, Data quality, Interoperability and Integratability. It identifies the need for clear policy mechanisms, champions of the concept, better understanding of governance and stronger ties with wider technological frameworks. One example from Europe which may exemplify this is the INSPIRE (Infrastructure for Spatial Information in Europe) initiative in the EU. This is at the stage of ratification and the result will be an agreed and fully developed model for public sector SDI across the member states that is to be implemented through an EU Directive in 2007.

At the same time it is important to be aware of realisable targets worldwide. One of the potential dangers of SDI, especially

to its proponents and some members of the GI community, is its aspirational aspect. The existence of strong legislative backing, which initiatives such as INSPIRE promote, may not be practical in other parts of the world. Progress may be slower in places which are either less technologically developed or where national interests around data protection and national security (such as in the US and China) may hamper that development for the moment. The unpredictability of human response and take-up are also issues that will continue to hamper development. It is ultimately important to point out that, quoting the case of the US NSDI;

*'the NSDI is not a concrete 'thing' but is more of a vision, a state of mind, a campaign and an enabler for better use of scarce resources'*  
(Source: Longley et. al. 2005).

### Further Reading

Burrough, P. and Masser, I. (Eds.). (1998) *European Geographic Information Infrastructures: Opportunities and pitfalls*, GISDATA 5, Masser, I. and Salgé F. (Eds.) London, Taylor & Francis Ltd. and European Science Foundation.

Cho, G. (2005) *Geographical Information Science: Mastering the Legal Issues*. Chichester, J. Wiley & Sons.

Crompvoets, J., Rajabifard, A., Bregt, A., Williamson, I. (2004). Assessing the worldwide developments of national spatial data clearinghouses. *International Journal of Geographical Information Science*, 18(7), 665-689.

Curry, M. R. (1998) *Digital Places: living with geographic information technologies*. London, Routledge.

Department of the Environment (1987) *Handling Geographical Information: Report of the Inquiry Chaired by Lord Chorley*. London, HMSO.

Federal Geographic Data Committee (1994) Executive Order 12906, <http://www.fgdc.gov/publications/docume>

[nts/geninfo/execord.html](http://www.fgdc.gov/geninfo/execord.html), Accessed March 10<sup>th</sup>, 2006.

Groot, R. and McLaughlin, J. (Eds.) (2000) *Geospatial data infrastructures: Concepts, cases and good practice*. Oxford, Oxford University Press.

Longhorn, R. and O'Riain, G. (2005) *Investigation of Best Practice Principles for a Spatial Data Infrastructure in the Technical Field*. Dublin, Report to the Irish Spatial Data Infrastructure Committee. November 2005.

Longley, P., Goodchild, M., Maguire, D. and Rhind, D. (2005) *Geographic Information Systems and Science*. 2<sup>nd</sup> Edition. Chichester, J. Wiley & Sons.

Martin, D. (1996) *Geographic Information Systems: socioeconomic applications*. 2<sup>nd</sup> Edition. London, Routledge.

Masser, I. (2005) *GIS Worlds: Creating Spatial Data Infrastructures*. Redlands, ESRI Press.

Nebert, D. (Ed.) (2004) *Developing spatial data infrastructures: The SDI Cookbook, Version 2.0*. Reston, GSDI.

Pickles, J. (Ed.) (1995) *Ground Truth: The Social Implications of Geographical Information Systems*. New York, Guilford Press.

Schuurman, N. (2005) *GIS: A Short Introduction*. Oxford, Blackwell.

Williamson, I, Rajabifard, A. and Feeney, M-E.F. (2003) *Developing SDI: From Concept to Reality*. London, Taylor and Francis.

---