The emergence of geoportals and their role in spatial data infrastructures

David J. Maguire a,*, Paul A. Longley b

a ESRI, 380 New York Street, Redlands, CA 92373, USA
b Department of Geography and Centre for Advanced Spatial Analysis, University College London, 1-19 Torrington Place, London WC1E 7HB, UK

Accepted in revised form 28 May 2004

Abstract

Geoportals are World Wide Web gateways that organize content and services such as directories, search tools, community information, support resources, data and applications. This paper traces the emergence of geoportals, outlining the significance of developments in enterprise GIS and national spatial data infrastructures (SDIs), with particular reference to the US experience. Our objectives are principally pedagogic, in order to relate the development of geoportals to SDI initiatives and to review recent technological breakthroughs—specifically the development of direct access facilities for application services and metadata records, and the facility to utilize services directly from conventional desktop GIS applications. We also discuss the contributions that geoportals and SDI have made to simplifying access to GI, and their contribution to diffusing GI concepts, databases, techniques and models. Finally, the role of geoportals in electronic government (e-Government) is considered.

Keywords: Geoportals; Spatial data infrastructure (SDI); NSDI; GIS; e-Gov

* Corresponding author.
E-mail addresses: dmaguire@esri.com (D.J. Maguire), p.longley@geog.ucl.ac.uk (P.A. Longley).
1. Introduction

In recent years geographic information (GI) has become increasingly important as a framework for describing patterns on the surface of the Earth and vast databases have been created to provide an inventory of natural and cultural resources. At the same time geographic analysis and modeling tools and techniques have been refined to help explain and predict contemporary and future patterns and processes. GI systems (GIS) are the very heart of this movement because they provide the digital tools necessary to implement the concepts of geographic thinking. Evolving in close juxtaposition, GI Science is providing the fundamental underpinnings and body of theory that is a necessary concomitant for progress in managing our interactions with the world (Longley, Goodchild, Maguire, & Rhind, 2001).

From its early origins as a desktop, project-oriented technology GI systems have grown to become enterprise information systems (Bishr & Radwan, 2000; Coleman, 1999). Although it is not possible to define enterprise GI systems precisely because of their varied nature, some of the key attributes include: large numbers of users (hundreds to thousands) spread over multiple locations; large databases (hundreds of gigabytes, to terabytes); multiple, often complex and mission critical, applications; the use of commercial off-the-shelf industrial strength information system hardware and software products (including relational database management systems and web servers); and ongoing exacting requirements for project and system management (including resource and funding coordination).

Many enterprise GI systems originated ‘bottom up’ initially as intraorganizational initiatives driven by benefit:cost analysis and return on investment decisions about being able to perform existing tasks cheaper, better or faster. Others were developed to take on new tasks that would, amongst other things, generate income, or lead to greater customer satisfaction. Building enterprise systems of any type is fraught with pitfalls and GI systems are no exception. Coordination of parallel initiatives must reconcile different technology standards, administrative schema and funding regimes. Not surprisingly, given these and other problems, bottom up approaches to building interorganizational enterprise GIS have met with limited success. At the same time there has been a move to initiate top-down programs that establish infrastructure for all GI users, both small and large. These spatial (sometimes the term Geospatial is used interchangeably) data infrastructure (SDI) programs are increasingly viewed as the route to build capacity and encourage GIS participation. Geoportals—gateways to geographic content and capabilities—are a key element of SDI. This paper considers the development of geoportals and their significance to SDI.

2. Spatial data infrastructure (SDI)

In the 1980s many national surveying and mapping agencies felt motivated to initiate strategies for providing greater access to standardized GI (Groot & McLaughlin, 2000; Williamson, Rajabifard, & Feeney, 2003). The term spatial data
infrastructure (SDI) was coined in 1993 by the US National Research Council (Mapping Sciences Committee, 1993) to describe, amongst other things the provision of standardized GI access. The US Federal Geographic Data Committee (FGDC) defines SDI as the totality of ‘technology, policies, standards, human resources, and related activities necessary to acquire, process, distribute, use, maintain, and preserve spatial data throughout all levels of government, the private and non-profit sectors, and academia’ (www.fgdc.gov). Although FGDC define SDI cogently, Williamson et al. (2003) argue that the SDI concept ‘continues to evolve as it becomes a core infrastructure supporting economic development, environmental management and social stability in developed and developing countries alike. Due to its dynamic and complex nature it is still a fuzzy concept to many, with practitioners, researchers and governments adopting different perspectives depending on their needs and circumstances.’ SDIs exist at many scales from the global, to national, state, regional and local, though the same basic principles apply across all the scales. In the past decade more than 100 SDI programs have been established within and between many countries at local, regional, national and global scales (Craglia & Masser, 2002; http://www.spatial.maine.edu/~onsrud/GSDI.htm; GINIE, 2003; Lachman, Wong, Knopman, & Gavin, 2001; Lance, 2003; Masser, 1998; Van Loenen & Kok, 2004).

Effective use of GI requires easy access to documentation (in hard copy or more frequently digital form) that describes the provenance, ownership, quality, age, fitness for purpose and other useful properties. This ‘associated’ data documentation, or data about data is referred to as metadata. A key component of any SDI is a catalog of metadata that can be queried to search for data and resources using space, time and thematic attributes. Such is the importance of metadata that several standards have been developed to define key data descriptors, and GIS software systems include capabilities to capture, store and search metadata.

In the US the FGDC coordinates the implementation of the national SDI. The vision for NSDI is that ‘current and accurate geospatial data will be readily available to contribute locally, nationally, and globally to economic growth, environmental quality and stability, and social progress’ (FGDC, 1994). The US NSDI program is widely viewed as the archetypal example of an SDI because it is the most mature, extensive and well funded initiative (FGDC, 2004; Masser, 1998, 1999). It developed along three parallel fronts: a set of standards for describing, accessing and exchanging digital data; a clearinghouse network offering on-line access to metadata; and a set of framework data sets (e.g. administrative boundaries, orthophotography, and rivers) that cover the whole country (Longley, Goodchild, Maguire, & Rhind, 2001; Ryan, DeMulder, DeLoatch, Garie, & Sideralis, 2004). Without doubt, it has achieved its principal goals of spreading awareness, creating community involvement, building capacity, and establishing standards for accessing GI. Nevertheless, several lessons were learnt during the first phase of its operation. Although the program received backing at the highest political level in the form of an Executive Order signed by President Clinton in 1994 (Executive Order #12906), this Order only pertained to Federal agencies. It did not relate to other tiers of government, or to the private sector—both major participants in the GI community. Furthermore, the
ideas encapsulated in the Order were not backed by financial control because there were no budgetary ties. This made it easy for people to ignore the suggestions of collaboration to create an SDI. The Clinton Executive Order established the FGDC as the NSDI coordinator. This organization, under the control of the United States Geologic Survey (USGS), did much to germinate and grow the NSDI, but in retrospect its dominantly technical focus, and lack of attention to issues of governance and policy, stymied widespread acceptance across the wider constituency of federal, state and local government organizations, as well as those in the private sector. This asymmetry of approach has also limited the durability of some of the NSDI’s achievements. In the decade since the US NSDI started, the technology for data sharing has also advanced in leaps and bounds, especially in the areas of web services and direct data access, rendering obsolete the first generation systems such as the Clearinghouse Network (http://www.fgdc.gov/clearinghouse/clearinghouse.html; Rhind, 1999).

In 2002 a second generation of the US NSDI was established under the auspices of President Bush’s e-government program. The e-gov program, administered by the Office of Management and Budget (OMB) entails a management agenda of 24 projects that seek to make government more focused on citizens and on the generation of tangible outputs. Geospatial One-Stop (GOS) is an initiative that it part of this program. Consistent with the long term goals of the NSDI, GOS aims to promote coordination and alignment of geospatial data collection and maintenance among all levels of government (http://www.geo-one-stop.gov/). GOS builds on the first phase of the US NSDI by providing federal, state, and local governments, as well as private citizens, with ‘one-stop’ access to GI. As such, the goals of GOS are (FGDC, 2004): to establish a web-based portal for one-stop access to maps, data and other geospatial data and services (www.geodata.gov); to institute a collaborative process to develop data content standards ensuring consistency among data sets and allowing governments to share data and integrate multiple sources of information; to create an easy-to-access inventory of currently available data collected by Federal agencies; and to cultivate a planned data investment marketplace that will allow Federal, State and local governments to combine resources with one another on future data collection/purchase plans.

Although GOS is still in an early phase in its evolution, there are obvious signs of success in the form of clear community acceptance. In a typical week (April 2004) there were 5622 user visits to the site, and 503,709 page views with an average of 35,979 per day (http://www.geodata.gov/gos). As of April 14, 2004, 305 separate publishers had registered 9672 publications (metadata records) on the site. The major categories of access in the first six months were imagery base maps (15.5%), administrative and political boundaries (10.6%), and geological and geophysical data (10.2%).

Even though GOS is comparatively well developed it is by no means unique. There are similar initiatives in Europe, such as the EU wide INSPIRE (Infrastructure for Spatial Information in Europe) project (Annoni et al., 2004; Bernard, Kanellopoulos, Annoni, & Smits, this issue) and in Asia, such as the Australian SDI (Busby & Kelly, 2004), and Indian SDI (Sivakumar, Rao, & Dasgupta, 2004).
3. Geoportals

The foregoing discussion on SDI/NSDI describes the key developments that led to the need for geoportals and some of the functions they are required to perform. The word ‘portal’ stems from the Latin word porta and indicates an entrance point (Annoni et al., 2004). Portals are web sites that act as a door or gateway to a collection of information resources, including data sets, services, cookbooks, news, tutorials, tools and an organized collection of links to many other sites usually through catalogs. Thus a Portal is a web environment that allows an organization or a community of information users and providers to aggregate and share content and create consensus. There are general portals and specialized or niche portals (http://whatis.techtarget.com/definition/0,,sid9_gci212810,00.html: also see Fig. 1). Some major general portals include Yahoo, Excite, Netscape, Lycos, CNET, Microsoft Network, and America Online’s AOL.com. Examples of niche portals include Garden.com (for gardeners), Fool.com (for investors), and SearchNetworking.com (for network administrators).

Tait (this issue) defines a geoportal as ‘a web site considered to be an entry point to geographic content on the web or, more simply, a web site where geographic content can be discovered’. Already there are several examples of geoportals including the British Geological Survey (http://www.bgs.ac.uk/geoportal/home.html) that covers geoscience resources, US Geospatial One Stop (www.geodata.gov) and EU INSPIRE (http://eu-geoportal.jrc.it/) that deal with national government data, and the Geography Network (www.geographynetwork.com, and see Tait, this issue) and GSDI portal (http://gateway.gdsi.org/weswww/portal/index.html) that index a wide variety of global geographic data, to cite just a few. Geoportals organize content and services such as directories, search tools, community information, support resources, data and applications. They provide capabilities to query metadata records for relevant data and services, and then link directly to the on-line content services themselves. They can also control commercial usage of services by facilitating the sale/purchase of data and services.

Fig. 1. A classification of geoportals.
It is useful to subdivide geoportals into two groups (Fig. 1): catalog geoportals and application geoportals. Catalog geoportals are concerned primarily with organizing and managing access to GI; for example, GOS and The Geography Network. Application portals provide on-line, dynamic geographic web services; for example, Mapquest provides routing services (www.mapquest.com), National Geographic provides mapping services (http://www.nationalgeographic.com/maps/) and local and regional government portals support transport and planning portals for the UK (Beaumont, Longley, & Maguire, this issue). A prominent feature of all SDI geoportals is a catalog service for publishing and accessing metadata. The more advanced SDI programs are also beginning to feature application services.

Geoportals are built using underlying World Wide Web infrastructure technology and commercial off the shelf GIS software. Network communication between clients and web servers uses HTTP (Hypertext Transmission Protocol). Technically speaking, a geoportal is essentially a master web site, connected to a web server, which contains a database of metadata information about geographic data and services. The services are built and exposed as web services, that is, self-contained, self-describing web applications that can be invoked over the web using messages encoded in XML (eXtensible Markup Language) and transmitted over a HTTP connection.

A geoportal database is populated with metadata records of published GI Services (Operation 1 in Fig. 2). Users can issue queries against the database either from a lightweight web client, or a heavier-weight desktop GIS client, providing that they have an Internet connection (Operation 2). This allows users to discover what services are available on particular topic, geographic area, and time period combinations (Operation 3, illustrated in Fig. 3). The services can then be directly used in client applications (Operation 4).

Metadata describing geoservices can be loaded into a geoportal from publisher systems which are themselves GI systems (desktop-, server- or Internet-based). Geoportals provide search, e-commerce and administration capabilities and, as such, are the heart of SDI systems (Fig. 4). The user views the front-end content of the
geoportal as a web site with a collection of pages that describe content, search and navigation instructions, as well as information of general interest to the SDI community. The geoportal front end typically sits on top of an Internet Map Server (IMS) that delivers local services for metadata management, mapping, geocoding, data download and links to other remote sites. The data for the local geoservices is held in a Database Management System (DBMS) that is accessed via a Database Gateway that links the IMS with the DBMS. Queries to the portal can come from both thin and thick (desktop) clients over HTTP-based Internet connections. In this way, both professional GIS desktop users and casual browser-based users can make use of the portal and fully distributed services over an open Internet connection.

Two key technical breakthroughs distinguish the second generation geoportals from their earlier first generation clearinghouse counterparts. First, it is now possible to access directly both metadata records describing services and the actual services (mapping, data download, geocoding, routing, etc.) themselves. Secondly, services can be accessed from both conventional desktop GIS applications, as well as a thin client browser. When accessed from thicker clients, usage now includes support for combining data from multiple remote services over a web connection. For example, Fig. 5 shows up to date weather information (1 h rainfall from the Meteorlogix...
service) as a transparent overlay on top of a base map (state boundaries, roads, hydrology, land cover and elevation from the National Map in Reston Virginia).
While the technical perspective is very useful to understand what geoportals are and how they work, to truly understand their significance a second organizational/institutional perspective must also be considered. This perspective can be represented by examining the purpose of some geoportal initiatives. GOS (FGDC, 2004) was developed in order to facilitate intergovernmental and interagency partnerships that help leverage investments and reduce duplication; to stimulate collaborative approaches to sharing and stewardship of data; to highlight the value of geospatial information to support decision making and the business of government; and to build communities around data categories to serve as data stewardship leaders and teams to maintain currency of the portal. The availability of up-to-date and accessible information that can be gathered from many sources and compiled easily to use for multiple purposes helps leverage resources and support programs such as community planning for economic development, assessing environmental impact, improving disaster response and strengthening homeland security.

The EU equivalent of GOS, INSPIRE (Annoni et al., 2004; Bernard et al., this issue), aims to remove the five main obstacles which the authors perceive are holding back the development of the EU: gaps in spatial data including missing or incomplete data; a lack of documentation including metadata that limits the reuse of data; incompatibility of spatial datasets that restricted the ability to combine multiple datasets; incompatible geographic information systems that make it difficult to find, access and reuse spatial data; obstructions to sharing and reusing spatial data such as cultural, institutional, financial and legal barriers that prevent or delay the use of existing spatial data.

4. Implications of geoportals

The technical breakthroughs described above, when coupled with improvements in GI standards, institutional arrangements, delivery and access mechanisms, have led to profound changes in the use of GI and in the policy formation process of both government and private organizations. Easy access can now be provided to up to date information that is always available (subject to Internet access) and the publishers of data can encourage proper use of information by providing metadata on provenance, quality, age and fitness for purpose. Technology is now the least pressing problem as far as policy makers are concerned. The focus has shifted instead to GI legal, economic, and social issues (Van Loenen & Kok, 2004; Williamson et al., 2003).

Geoportals and SDI have made a major contribution to simplifying access to GI, and in so doing have helped to encourage and assist people who want to use GI concepts, databases, techniques and models in their work. Today geoportals are being built using industrial strength, industry standard commercial off the shelf information system technology. Geoportals can connect widely distributed GI providers and users via the medium of the Internet. Catalog geoportals facilitate GI sharing, avoid duplication and offer up to date geoservices with associated metadata about quality and fitness for purpose. The services to which they provide access can add
value to raw data by encapsulating data and tools in a single user-oriented application. Application geoportals provide advanced GIS functionality and can aggregate sub-services irrespective of location. When geoportals are used in conjunction with powerful mobile, desktop or server GIS packages they offer compelling capabilities for a wide range of GIS applications, ranging from facilities management, to the processing of planning or permitting applications, and journey planning. Although considerable progress has been made in the past three years, it is clear that this is only the start of a long and fruitful journey that promises a lot for users across government, academia and the commercial sectors. It is to be expected that existing geoportals will hold both more and better records about available GI datasets and that new application geoportals will offer an ever increasing range of services such as flood plain mapping, site suitability analysis, land use planning, travel and routing.

SDI programs are really part of a much wider trend of electronically-enabling organizations. Many governments throughout the world (Curtain, Sommer, & Vis-Sommer, 2004) have invested heavily in electronic government (e-gov) programs. Kim (2003) defines e-gov initiatives as ‘citizen-centered information services for whomever, whenever and wherever’. According to Song (2003) some of the key benefits of e-gov programs include: improved decisions about IT systems investment; alignment of IT support and business objectives/drivers; reduced redundancy; improved interoperability between processes and systems; realizing economies of scale; building services around citizens; making governments more accessible; social inclusion; and better use of information. Song (2003) details five stages of e-gov maturity beginning with (1) government-centered static information, progressing to (2) dynamic information spread over many separate site, to (3) interactive information (such as the ability to download forms), to (4) on-line transactions, and (5) ending with seamless integration of services across departmental and administrative boundaries. Second generation geoportals, like GOS and INSPIRE, have elements of the fourth and fifth stages.

SDI programs have attracted considerable interest and resources because their government sponsors not only recognize the strategic value of GI as a foundation stone for good governance, but often also because they are a potential vehicle for reducing costs and generating revenue to supplement other more convention streams of income (Song, 2003). SDI programs allow government activities to be transparent, and can increase public participation in the democratic process (through, for example, on-line voting). Establishing ready access to up-to-date GI provides a means of empowering government workers and citizens, often through channels for which that they are (directly or indirectly) prepared to pay. Improved access to higher quality information can improve decision making both within government and by external agencies. Moreover, the ability to collect data once and use it many times, whilst at the same time avoiding development of duplicate data sets and reducing data management costs, also presents a cost effective way of widening access to information resources (Blakemore, 2003; Rhind, 1999).

On their own geoportals are not, of course, a universal panacea for all that is wrong in government and commercial organizations. Geoportals can serve to emphasize the digital and financial divide evident in many societies (Kim, 2003).
Issues of security (Baker et al., 2004) and privacy are even more relevant and acute when there is greater access to information and government activities are opened up to wider scrutiny. Finally, there are resource challenges to be addressed whenever existing practices and processes are re-engineered.

The remaining contributions to this theme issue illustrate the state-of-the-art of portal development through a range of applications: Tait (this volume) provides an extended definition of the geoportals concept, and illustrates this with four commercial and government applications from the US; Askew, Evans, Matthews, and Swanton (this issue) describes and evaluates an extended example of partnership of UK government departments and agencies that is providing access to rural and countryside information; Beaumont et al. (this issue) appraise the range of applications that are developing across the UK government sector at the national, regional and local levels; and Bernard et al. (this volume) address the problems and prospects for development of a pan-European SDI. Together, these papers present a wide-ranging review of the central issues surrounding geoportals development and information exchange.

References


