

Representations of Environmental Data in Web-based GIS

Peter Mooney and Adam C. Winstanley,
Department of Computer Science,
National Center for Geocomputation,
National University of Ireland Maynooth,
Co. Kildare, Ireland
Tel: +353 1 268 0100
Fax: + 353 1 268 0199
Email: {peter.mooney¹, adam.winstanley}@nuim.ie

Abstract

The GIS community is using the vast potential of the Internet to disseminate geospatial information. Web-based GIS software and services are key components in distribution of geospatial data. Web-based GIS provide government departments, local authorities and environmental agencies with unprecedented opportunities to offer online access to their environmental information and related services for citizens. Web-based GIS offers access to information services 24 hours a day, 7 days a week, 365 days a year. In order for web-GIS to be successful in delivering environmental information the representation of the input datasets and output delivery formats/structures must be suitable to both the Internet delivery medium and the intended audience. In the majority of cases this will involve conversion and re-modelling of existing data resources. This paper discusses representations of environmental data for delivery and dissemination using web-based GIS in order to serve a variety of stakeholders : policy makers, scientists, media, and the general public. We summarise the major issues for delivering complex geospatial data about the environment using this medium. Prioritisation of metadata collection and geospatial data interoperability are crucial factors in delivering effective web-GIS tools. The INSPIRE Directive will greatly increase the number of available data sources and the use of web-based GIS for environmental information provision in the future will be discussed.

1 Introduction

Environmental issues are now topics of conversation for the general public. Coffee-break conversations often include global environmental issues such as climate change or changes in weather patterns and more localised issues such as air quality, water quality, and waste disposal. Previously environmental data access and distribution was confined to policy makers, analysts, and the environmental science community. Traditionally, GIS has been a technology which has run on large computer systems, then migrating to desktops and, more recently, the software has been able to work across the Internet. This has provided many opportunities to provide access to data previously unavailable to the public (Kingston, 2007). The Internet-enabled general public are now an important stakeholder for governments and environmental agencies across the world. Their awareness of environmental issues coupled with almost universal access to the Internet means that there is a great opportunity for web-based GIS to provide the services that stakeholders require. There is also an increasing recognition that government agencies holding environmental information collected at public expense must make these data assets accessible. This accessibility includes easier and quicker access to this environmental information for little or no charge.

The paper discusses the use of web-based GIS from the perspective of users who will use web-based GIS for either information gathering purposes (ie accessing local authority information) or to download geospatial data. In this paper we predominantly focus on users of web-based GIS. This user audience comprises of GIS specialists, decision makers, and the general public. If the user group has no formal GIS skills we refer to them as non-specialist users. The core discussion topic in the paper is how web-based GIS systems are developed from a software perspective, managed by organisations, and used to deliver visualisation and geospatial data access services for the

¹ Corresponding Author: This author is funded under a Postdoctoral Fellowship with the Environmental Research Center of the Irish Environmental Protection Agency.

users mentioned above. This involves a discussion of the most popular types of representations for environmental data being delivered using web-based GIS and the role metadata has to play as a representation of data and software services. Web-GIS present a unique opportunity in data and information provision to make more data available to a wider range of people.

The paper is organised as follows. Section 2 outlines an overview of web-based GIS by discussing the key components in delivering GIS to the Internet. This is followed by a brief overview of some of the most well known and widely deployed web-GIS server and client software technologies. Section 3 discussed the core focus of the paper by looking at how best to represent environmental information in order to allow it be used by a wide variety of web-based GIS applications and tools. Section 4 looks at metadata and its importance to web-based GIS and to access to geospatial data in general. The paper closes with a summary of the main conclusions of the paper and some ideas for further work and discussion.

2 An overview of web-based GIS

Web mapping is loosely defined as the process of designing, generating and delivering maps on the Internet. Web GIS brings GIS functionality to web mapping. In GIS vernacular web mapping and web-based GIS are often used synonymously despite the fact that they do not necessarily refer to equivalent technologies. During the mid nineties several innovations made it possible to develop web-GIS solutions: support for vector graphics in Internet browsers, the birth of Javascript and JAVA, ESRI's entry into web mapping amongst other technological developments. Such software development was very complex and was only carried out by large organisations with large software development budgets. For other users delivering map-based information on the Internet was facilitated in many cases by generating the necessary mapping as JPEG or GIF images and then implementing a mix of HTML and Javascript to allow users to click on features in these maps, query and retrieve other map output, or access data download. This situation has greatly changed and users of web-based GIS perform many of the operations they can perform on their desktop GIS. Examples of such operations include: pan and zoom, feature (point, line, polygon) based querying, addition or removal of layers, path distance calculations, area estimations, buffer zone queries, etc. In theory any spatial analysis or spatial statistics functionality available in standard desktop GIS can be implemented within web-based GIS. However there is one very important consideration. Desktop GIS benefit from the processing power of the local computer, have immediate access to disk storage, and high end graphics visualisation. In web-based GIS all operations must be performed in real-time. Factors such as bandwidth capacity, network latency and Internet browser type must be taken into consideration. In addition to this web users expect near to instantaneous responses from information systems (Waller and Sharpe, 2006). For this reason only lightweight queries (any query to which the user is delivered a response in close to real-time) are usually implemented. Large scale queries are performed off-line.

2.1 Web-based GIS Technologies

Web-GIS applications can be classified according to whether they are server or client systems. These shall be discussed in more detail in the next sections. Before this discussion we will outline the two key architectures upon which web-based GIS are developed. The first is the Open Geospatial Consortium specifications and the second is AJAX (Asynchronous Javascript and XML).

The Open Geospatial Consortium (OGC) is an international voluntary consensus standards organisation. The OGC specifications are a collection of specifications or standards developed to assist in achieving interoperable geospatial technology at both the software service level and the geospatial data specification level. The most widely implemented OGC specifications are as follows: Web Map Services (WMS) which provide specifications on how the delivery and rendering of maps as images should be implemented in software; Web Feature Services (WFS) specifies the standards for the exchange of raw geographic feature data over the Internet; Web Coverage Services (WCS) are similar to WFS are used for the exchange of geographic coverage data such as segmented curves, grids, and digital terrain models. The OGC Specifications have greatly influenced the direction of recent web-based GIS developments. WMS and WFS makes it much

easier to publish, visualise, and exchange any geospatial data over the Internet. WMS essentially creates maps in popular image formats (PNG – Portable Network Graphics, or JPEG formats) of the requested geographical area. The generated image is rendered within the web browser. Users do not have to copy large geospatial datasets to local systems in order to visualise the output in their desktop GIS. WFS allows users to access and download subsets of larger geospatial datasets directly into their web-based or desktop GIS.

AJAX is primarily an Internet application development technology which supports the development of web pages that permit interactivity without the need for web page refreshing. This is particularly useful when map-based interfaces are embedded into web-based applications. If the application is driven by AJAX the client user can query and interact with the map display without the need for web page refreshing. Many readers will be already familiar with this technology from using web sites featuring Google Maps or Microsoft Virtual Earth mapping.

If organisations do not wish to manage web-based GIS server or client software directly from their own systems Google Maps and Virtual Earth (for example) allow developers to access their API (Application Programming Interfaces) to develop customised location-based services and web mapping services. This means that such users do not have to manage web-based GIS services on their own systems and can take advantage of the large cartographical resources of Google and Microsoft. Sharl (2007) remarks that this approach by Google and Microsoft have exposed the deficiencies of traditional GIS by tapping into what GIS end users really want: simplicity, accessibility, immediacy, responsiveness, and low cost development of web mapping services. The use of AJAX technology has shown organisations and users that access to geospatial information and build web mapping services can be easy and intuitive.

Consequently general Internet users with little or no GIS, cartography, or geospatial data handling expertise are equipped with a very powerful set of mapping tools and services. Such is the simplicity of using the Google Maps or Microsoft Virtual Earth APIs that in just a few minutes users can integrate mapping into their websites or blogs by simply copying and pasting embeddable code. The very simple representation of geospatial information using KML (Keyhole Markup Language – see section 3.3) has empowered a vast community outside of GIS. Sharl (2007) remarks that this opens a new age of cartography that people are now “spurred by space photography, mobile phones, and new ways of annotating Web content that the ancient art of cartography is now on the cutting edge.”

2.2 Key Web-based GIS Server Software Tools

Geospatial data is usually stored on a server machine in order to make it access for Internet access. In order to provide spatial query functionality, cartographic visualisations, and spatial data exchange special GIS software must be running on these server machines. This software is referred to as GIS Server Software. This section gives a brief overview of four of the most popular GIS server software tools.

ESRI ArcIMS. ArcIMS (Arc Internet Map Server) is a web map server developed by ESRI. Web-mapping applications developed on ArcIMS offer tools users of other ESRI products will be familiar with – zooming, toggling between layers, identify and query, find and measure, buffering, select by feature, and printing of the cartographic output. ArcIMS can be customised and extended using programming environments such as JAVA, ASP, and .NET. It is particularly well suited to organisations and individuals who already use other ESRI products (ArcSDE and Geodatabases) to manage their geospatial data resources.

MapGuide Open Source: MapGuide is an open source version of Autodesk's MapGuide Enterprise. MapGuide is a web-GIS server allowing developers to develop interactive web mapping applications by using AJAX. It offers access to most vector and raster data formats including SDF vector file formats and CAD-based data access. MapGuide offers quality cartographic output, uniform access to raster and vector data formats, application development in several languages such as PHP, .NET, and JAVA, and access to mapping and feature data from other publicly available WMS and WFS on the Internet.

MapServer: MapServer is also an open source development environment. It delivers high quality

cartographic rendering of spatial data (maps and images) and access to spatial data (vector and raster). It also features a fully featured development environment allowing developers to create applications to interact with MapServer using environments such as PHP, Ruby-on-Rails, C#, and JAVA. Many organisations use MapServer as the access point to their geospatial data repositories to take advantage of the excellent WMS, WFS, and WCS implementation. Consequently MapServer is often used in conjunction with a dedicated web-based GIS mapping client where these clients render the output from MapServer in client browsers and applications.

GeoServer: GeoServer is a JAVA driven geospatial data server providing organisations with WMS, WFS and WCS implementations. In addition to vector and raster support it offers mature and stable support for a large variety of geospatial databases such as PostGIS, Shapefiles, ArcSDE, and Oracle. As it is a JAVA servlet based application it can be run within any java servlet container environment. A large deal of the functionality offered by GeoServer ie reprojection, access to geospatial databases, raster manipulation, are handled transparently by GeoServer itself. It uses the extensive JAVA GIS toolkit called GeoTools to accomplish this.

2.3 Web-based GIS Client Software Tools

There are a large number of software tools available which offer easy-to-use client-side web-mapping application development and access to geospatial data services. This section provides an overview of three of the most commonly used tools. Web-based GIS clients normally connect to publicly accessible WMS, WFS, and WCS offered by data provider organisations running web-based GIS server technologies such as those mentioned above in Section 2.2. WMS are also offered by Google (through Google Earth and Google Maps) and Microsoft (through Windows Live and Virtual Earth). *ESRI ArcIMS* features a rich set of web publishing capabilities allowing organisations to develop scalable and interactive interfaces to their geospatial data resources. The key advantage of using ArcIMS as the development environment is that the overall cost-of-ownership is controlled as ArcIMS also controls the Web-GIS server functionality and management. *GeoMedia WebMap* by Intergraph is a long established web-based map visualisation and analysis software tool. It can access and view data a large number of geospatial data formats in their native format without conversion or translation. Additionally has extensive Scalable Vector Graphics (SGV) support, which is a World Wide Web Consortium (W3C) graphics format specification. The Free and Open Source *OpenLayers* project is a pure Javascript library for displaying map data in web browsers with no server-side dependencies. The OpenLayers Javascript API provides developers with full control over OpenLayers-powered maps from within Javascript on a web page. OpenLayers offers no server capabilities but can connect to a wide variety of WMS, WFS, and WCS.

Figure 1 is an example of a web-based GIS client accessing remote GIS services and rendering a map within the web-browser. The example uses OpenLayers API to access a OGC WMS at University of Leeds UK and a KML file which is available at the Environmental Protection Agency Ireland. The Javascript API hides all complex details of accessing WMS or other data resources and renders the map in the browser quickly and easily. The rendered map is shown below in Figure 2. Readers familiar with development of web-GIS client interfaces in Google Maps and/or Microsoft Virtual Earth will notice the similarities between these two Javascript APIs and the OpenLayers Javascript API.

```
<script type="text/javascript">
var map, layer;
function init(){
    map = new OpenLayers.Map( 'map' );
    var layer = new OpenLayers.Layer.WMS( "Global MODIS Blue Marble",
    http://iced.s.ge.ucl.ac.uk/cgi-bin/icedswms?", {layers: 'blumarble_1' } );
    map.addLayer(layer);
    map.addLayer(new OpenLayers.Layer.GML("NOAA Bouy Locations",
    "http://erc.epa.ie/kml/bouys.kml", {transparent: 'true', format: OpenLayers.Format.KML}));
}
```

```

map.addControl( new OpenLayers.Control.LayerSwitcher() );
map.setCenter(new OpenLayers.LonLat(-92.6835, 32.395), 5);
}
</script>

```

Figure 1. A simple example of using the OpenLayers Javascript API to access a WMS data source and a KML file.

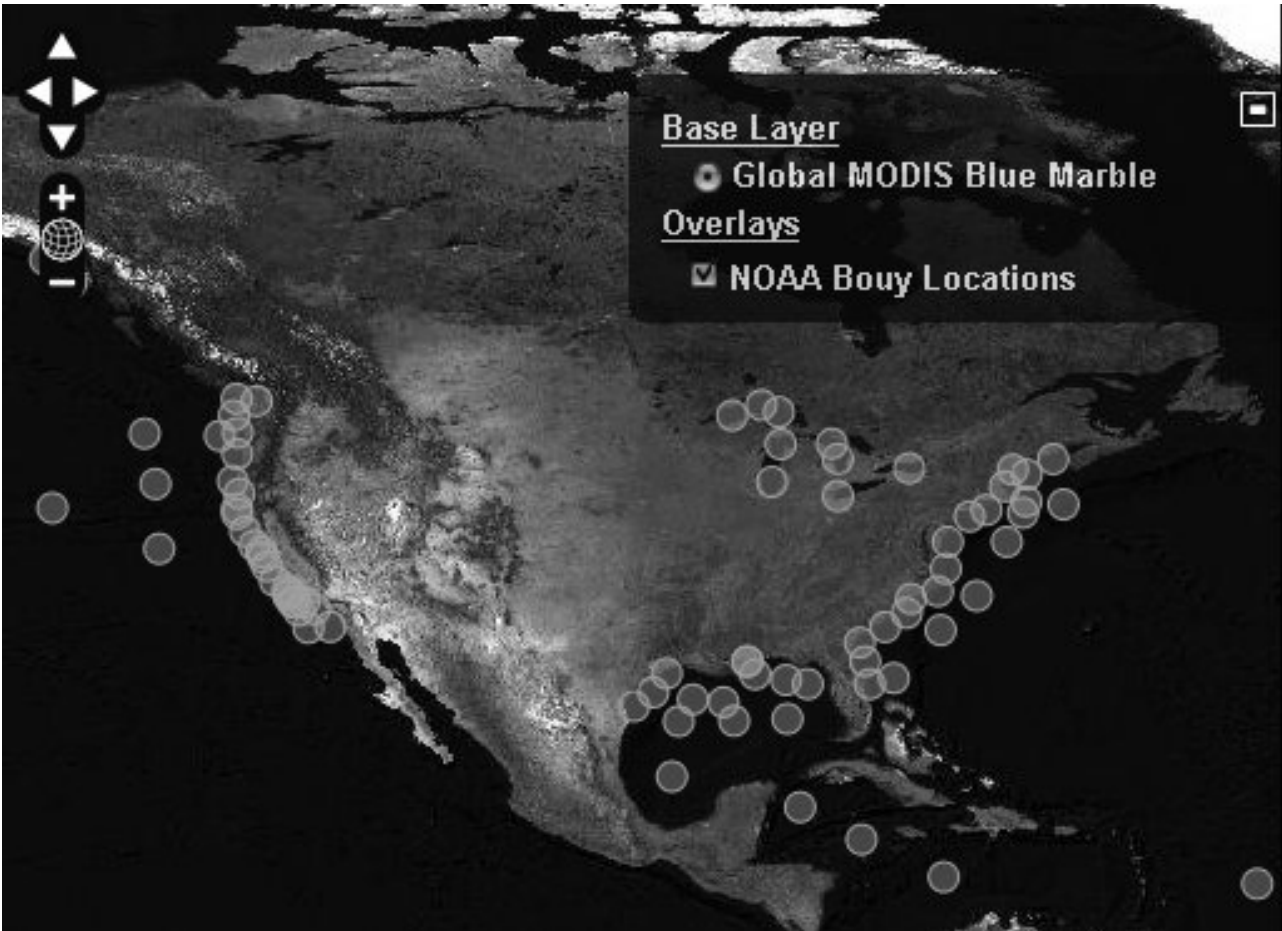


Figure 2: The cartographical output from the Javascript function in Figure 1. The screenshot features the pan & zoom and layer switcher functionality offered by the web-based GIS client.

3 Representing Environmental Data for web-based GIS

The tools discussed in Section 2 offer remarkable flexibility in terms of the wide range of geographical data formats and representations they can both import for analysis and visualisation and explor. The underlying principle for all of these tools is that the geospatial data is represented in standardised formats with several prominent standard formats emerging: ESRI Shapefiles , MapInfo TAB format, ArcSDE, etc. Choosing the correct format to store environmental data in is crucial in order to prevent duplication of work and waste of resources at a later stage converting to other formats or fixing problems in the current format. Mooney and Winstanley (2007b) comment that large amounts of geospatial data from the environmental domain is stored and manipulated in software storage formats which are unsuitable for the purpose. Incorrect choice of format often makes analysis of the dataset difficult while at the same time diminishing the opportunities for data re-use by other scientists. An example of this is using MS Excel to store time series, location-based, environmental monitoring data. Performing advanced location-based statistical queries on these datasets is difficult as MS Excel does not have the range of tools which can natively perform such analysis. The next sections outline the key features of three methods of representing

environmental data which allow better integration with web-based GIS systems.

3.1 PostGIS – A Spatially Enabled Database

A spatial database is a database which is optimised to store and query data related to objects in geographic space – points, lines, and polygons. PostGIS is a spatial language extension to the well known PostgreSQL database server. It implements the OGC Simple Feature Specifications for SQL standard which defines the operations required to insert, query, manipulate and manage spatial data objects. ESRI Shapefiles can be loaded directly into PostGIS using the internal PostGIS Shape Loader tool. Flat files of coordinates and attributes may also be loaded directly into the database with PostGIS performing the appropriate coordinate system conversions if necessary.

3.2 Geographic Markup Language (GML)

GML was developed by the Open GIS Consortium and is a spatially enabled dialect of XML. GML provides both a vendor neutral and a software implementation neutral format that is optimally suited for distribution over a network. Both plain GML files as well as compressed GML files may be streamed so that a user does not have to wait for downloading an entire file before opening; this greatly enhances usability in a networked environment. GML 2 (and now GML 3; recently approved by the OGC) can enabling linking of features in one GML file to those in another GML file through use of hyperlinks.

In Figure 3 below a small sample of GML is shown from the Ordnance Survey UK's OS MasterMap (Basingstoke Free Sample Dataset). The feature described in a line feature which is obstructing other features. It is visually easy to understand the attributes of this feature (its feature ID, the spatial themes it is classified under etc). The GML structure (with the assistance of the appropriate GML Schemas) parsing and extracting from GML data files is reasonably straightforward. As GML is an OGC standard it is one of the default import and export representation in many GIS software tools. All of the tools mentioned in Section 2 import GML natively and can display the contents of the dataset without conversation to an intermediate structure or format.

```
<osgb:TopographicLine fid='osgb1000000334391475'>
<osgb:featureCode>10046</osgb:featureCode><osgb:version>1</osgb:version>
<osgb:versionDate>2001-11-07</osgb:versionDate>
<osgb:theme>Land</osgb:theme><osgb:theme>Rail</osgb:theme>
<osgb:accuracyOfPosition>1.0m</osgb:accuracyOfPosition>
<osgb:changeHistory>
<osgb:changeDate>1970-01-01</osgb:changeDate>
<osgb:reasonForChange>New</osgb:reasonForChange>
</osgb:changeHistory>
<osgb:descriptiveGroup>General Feature</osgb:descriptiveGroup>
<osgb:physicalLevel>50</osgb:physicalLevel>
<osgb:physicalPresence>Obstructing</osgb:physicalPresence>
<osgb:polyline>
<gml:LineString srsName='osgb:BNG'>
<gml:coordinates>278235.950,187050.750      278231.100,187064.450      278227.250,187075.950
</gml:coordinates>
</gml:LineString></osgb:polyline>
</osgb:TopographicLine>
```

Figure 3: An example of a feature represented in GML from the Ordnance Survey UK OS MasterMap dataset.

Figure 4 shows a segment of GML for a point-based dataset. The dataset represents the location

of ground based air quality monitoring stations under the control of the Irish Environmental Protection Agency. The (easting, northing) of each station is represented. Other parameters include the type of aerosols monitored at that station. The spatial characteristics of the station are represented as standard within a geographical layer. The URL of a data processing service is provided. If a user queries this location from a web browser-based map display they are provided with the link to data access service. This service provides visualisations of the previous year's air quality monitoring data and access to the raw data in several different formats.

```
<ogr:AirStations fid = "1002">
<ogr:geometryProperty>
  <gml:Point>
    <gml:coordinates>253870.2,206530.8</gml:coordinates>
  </gml:Point>
<ogr:Site>Emo Court (Co. Laois)</ogr:Site>
<ogr:Type>Fixed Station</ogr:Type>
<ogr:LocationType>Rural</ogr:LocationType>
<ogr:Operator>EPA</ogr:Operator>
<ogr:Parameters>Ozone,SO2</ogr:Parameters>
<ogr:DLink>http://erc.epa.ie/air/query.jsp?location=EmoCourt</ogr:DLink>
</ogr:geometryProperty>
</ogr:AirStations>
```

Figure 4: An example of a GML representation of a point feature. The coordinates of the point feature is specified accompanied by some attribute information. The `<ogr:DLink>` attribute allows users to be redirected to a related service where more detailed data can be obtained corresponding to this feature.

3.3 Keyhole Markup Language (KML)

The key building block behind the types of applications we see on Google Maps and Google Earth is the KML language. KML is a XML variant. The entire schema is available from the Google website. KML allows users to represent their geospatial information in such a way that it can be overlaid onto the base-maps within Google Earth or Google Maps. The word "mashup" is the now accepted technical term for when a geographic layer is created in this manner. Geographical features such as points, lines, and polygons are expressed in KML as ordered list of (latitude, longitude) pairs (expressed in decimal degrees). Primarily for visualisation purposes, the popular uptake of mashups and Google Earth and Google Maps has seen KML develop into something of a defacto standard for exchange of datasets with reasonably non-complex sets of geographical features (points, lines, polygons). KML has recently been submitted to the Open Geospatial Consortium with a view to standardising its usage for geobrowser applications. Many of the tools mentioned in Section 2 can import KML datasets and display these datasets as base layers or overlays. Export to KML is also available in many cases. Due to the simplicity in representation of geographic features with KML it is relatively straightforward to write converter software which extracts output from a geospatial dataset and outputs the entire dataset (or selected subset) into a KML representation. With a short investment of time in learning how to create KML non-specialist GIS users can begin creating mashups using GPS data they have collected themselves. As GPS technology becomes more ubiquitous these type of data streams flow from a diverse and unconstrained set of sources: athletes tracking their training runs, birdwatchers following transects and sightings of birds, hikers mapping out paths in woods and forestry.

Figure 5 illustrates how the software in section 2 and the format, storage, and representation options in this section are integrated by organisations in building a spatial data infrastructure. At the backend, geospatial data is stored in a mixture of spatially enabled databases, relational databases, and file systems. Web-based GIS server software can then access these data sources using connections (such as JDBC for JAVA-based applications) or natively (such as ArcIMS with ESRI ShapeFiles). A rich client interface can be offered using web-based GIS client software. End-

users can retrieve geospatial data as images, feature data (in GML format), as KML, or be directed to other services such as in Figure 4.

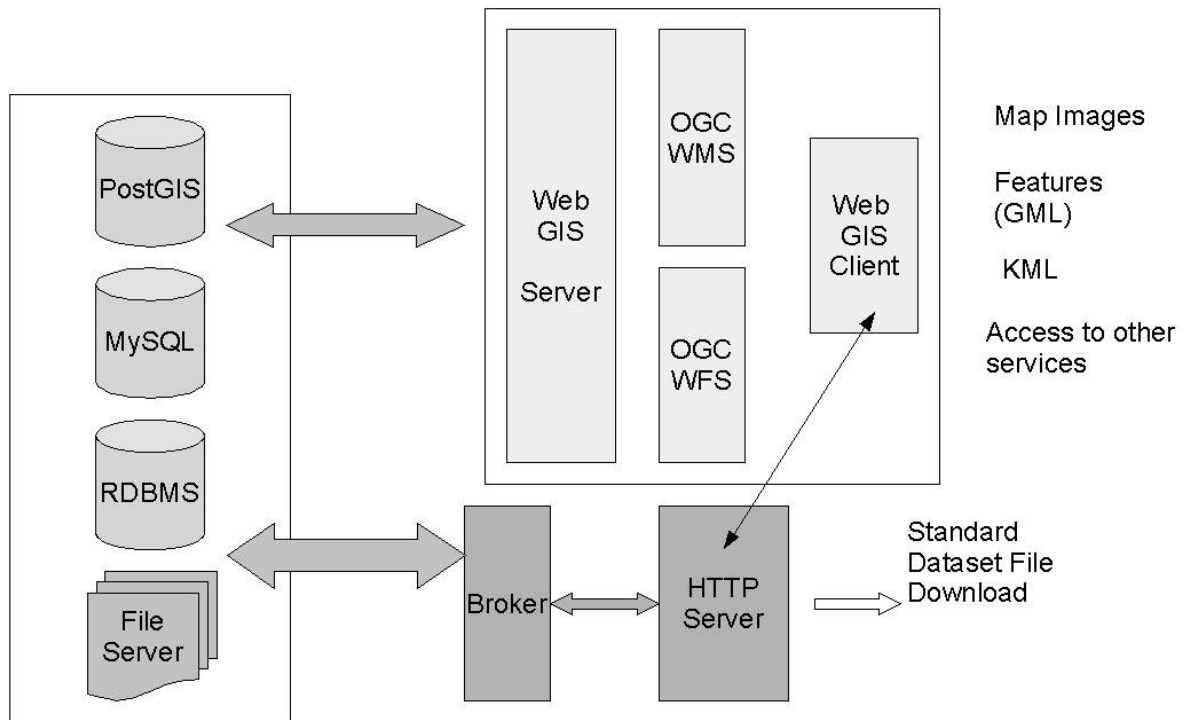


Figure 5: A schematic of the possible configuration of a spatial data infrastructure offering both GIS server and client functionality.

4 The role of metadata in web-based GIS

The discussion in the previous sections has focussed on software issues related to access to geospatial data. Numerous studies have established that while the value of geospatial data is recognised by government, the scientific communities, and society (Nerbert, 2004, Ch:3) the effective use of geospatial data in web-based GIS is inhibited by poor knowledge of the existence of these data, poorly documented information about the datasets, and data inconsistencies. The loss of time and resources in searching for existing spatial data or establishing whether they may be used for a particular purpose is a key obstacle to the full exploitation of the available data (INSPIRE, 2007). Metadata is an important resource in its own right. It provides a high level representation of the contents, physical properties, and geographical characteristics of the dataset. In the case of some GIS tools metadata can be consumed as a data resource if the representation is a recognised metadata standard such as ISO 19115. Metadata has become a term conveniently ignored or avoided by those required to provide or manage data. Consequently large collections of environmental and other geospatial data become “data tombs” (Han & Kamber, 2006) seldom visited or maintained, key datasets never emerging from “grey dusty archives” or “slowly rotting” (EteMII, 2004) because users do not know of their existence or are denied access to them. Metadata is seen as “boring” (Comber et al, 2005). Consequently many researchers see little or no academic or workplace recognition for the task of providing or maintaining metadata (Mooney & Winstanley, 2007b). As the volumes of geospatial data being generated increases, particularly in the area of environmental monitoring and environmental science, this situation must not be allowed continue. Some authors remark that it is ironic, that at a period where the volume of data

generated in scientific research is at an all time high, the practice of documenting and accessing these resources is at an all time low (Bulterman, 2004).

The concept of metadata was introduced to provide orientation in a space of continuously growing data and information resources (Pillmann et al, 2006). Metadata provides information about the data but does not include the data itself. In the majority of cases the metadata accompanying a geospatial data resource is represented within a tabular structure in a separate digital file. It may also be represented as rows within a database table (or set of related tables). Some earth science disciplines (climatology, weather forecasting) use machine-independent data formats that support the creation, access, and sharing of array-represented scientific data. Examples of these data formats include NetCDF (Network Common Data Form) and HDF (Hierarchical Data Format). These formats are self-describing, allowing a software application to interpret the structure, contents, and representation of the dataset file without any outside information. The metadata is embedded within the file format. For web-GIS metadata is usually stored in files with XML format or stored within tables of relational databases and converted to XML. Using style sheets the XML representation of the metadata is presented in a clear human readable representation (usually in conjunction with HTML) in the web-GIS.

4.1 Metadata Representations

For Web-GIS metadata can actually be considered as a data resource. When search tools are provided through the Web-GIS interface it is from the metadata resources where search results are extracted. The metadata then usually provides links to data visualisation and data extraction services. Without high-level representation through metadata the datasets themselves are essentially standalone digital objects. The metadata provides a explanation of what the data resources represent, what the map layers available in the web-GIS actually represent, etc. Waller and Sharpe (2006) remark that the importance of metadata cannot be overestimated. "It adds a whole new dimension, providing extra richness of contextual or descriptive information at the point of access". The perception amongst many scientists and creators/maintainers of geospatial data is that the creation and maintenance of geospatial metadata is a laborious and unnecessary data management task. Many GIS software packages (for example ESRI ArcCatalog) will automatically create metadata in a standardised format such as ISO 19115 as required by INSPIRE. Many of the fields are filled in automatically. These fields include: information about the parties responsible for the creation, analysis, and maintenance of the dataset resources; file types, sizes, file content, and other computer representations; and in the case of GIS formats (such as Shape Files) the geographical extent, scale, projections used. These information fields are vital in metadata provision. However, it is often the case that fields requiring manual input from the scientific expert (such as fields related to the quality or pre-processing of the dataset) are often left blank or only partially filled in. As a result it is very difficult for a third party to make an assessment of the fitness-for-use (quality) and the fitness-for-purpose (the problem they are trying to solve). This usually causes such third parties to download the partial or entire dataset into their GIS or analysis software in order for them to make these preliminary assessments. Metadata is also a core component of OGC Services. Figure 6 shows a subset of the metadata returned from a GetCapabilities request from a WMS. In this example the information contained in the metadata can be used automatically by the web-based GIS client to perform the necessary coordinate transformations.

```
- <Layer queryable="0" opaque="0" noSubsets="0">
  <Title>GDR_E</Title>
<SRS>EPSG:4326</SRS><SRS>EPSG:4269</SRS><SRS>EPSG:4267</SRS>
  <LatLonBoundingBox minx="-150" miny="40" maxx="-47" maxy="90" />
  <BoundingBox SRS="EPSG:4269" minx="-123.6486" miny="48.8696555569"
  maxx="-123.0846040092" maxy="49.1189004763" />
- <Layer queryable="0" opaque="0" noSubsets="0">
  <Name>AtlanticDEM</Name>
```

```
<Title>Atlantic Canada Digital Elevation Model</Title>
<SRS>EPSG:4326</SRS>
<LatLonBoundingBox minx="-72.000849" miny="39.996615" maxx="-48.002030" maxy="51.997853920" />
</Layer>
```

Figure 6: Subset of the XML from a `GetCapabilities()` call to the WMS at National Resources Canada. The XML shows the base layer `GDR_E` with the layer `AtlanticDEM` as a sub-layer. The SRS denotes the coordinate system transformations allowed on these layers.

5 Conclusions

There are many advantages to providing GIS services on the Internet and WWW. Many mapping and visualisation services are being made available using Web-based GIS systems. Stachowicz (2006) summarises the most important advantages for users accessing these services as (A) that there are no software downloads required as usually an up-to-date web-browser is all that is required; (B) users are offered one interface to potentially many separate services – for example geospatial data integrated from an environmental organisation, human health research, and land-user planning; (C) users can access these services in a 24-7-365 manner. NSB (2005) argues that much of the scientific data being collected today are “born digital”. There is no analogue or paper counterpart. Additional scientific data are being converted to digital representations and being disassociated with their analogue representations. For these data resources the public cannot merely make an appointment with a local authority office and physically visit to browse the files. Web-based GIS allows authorities to provide “always on” services. Where appropriate users can browse map-based representations of these datasets and possibly download the data to their own computer for further analysis.

5.1 Metadata and Improving Access to Geospatial Data

As the volume of geospatial data about the environment continues to grow so to does the need to properly document these data resources with metadata. Without metadata these data resources may lie dormant and undiscovered on the Internet. This invisibility may give rise to duplication of effort in creating, accessing, and managing these data resources. The INSPIRE directive states that geospatial data must be managed as close to the source as possible. In relation to this the representation of key knowledge about these data resources in metadata must be performed initially as close to the original data creator source or scientific expert group. Without this vital specialist knowledge about the data resources may be left in “dusty archives and grey literature” (Wilson, 2002) or lost in laboratory or field notebooks. The management overhead of convincing busy scientists and analytical staff that they need to maintain and manage metadata for their data resources is considerable. Exposing geospatial data through web-services for consumption by a wide variety of users, who may be using web-GIS to access these services, is greatly hindered if the users do not have access to accurate and complete metadata.

The authors feel that an important feature of web-based GIS is that their usage can indirectly assist in the efforts of digital resource curation and ensuring long-term access to geospatial data resources. Many scientific funding organisations highlight urgent needs to invest in data curation and data recovery. “Very substantial amounts of data have already been lost and even greater losses are imminent as the ability to recover data stored on obsolete technologies declines exponentially” (Leach et al, 2004). Web-GIS systems are encouraged to access the underlying geospatial data in its original format (or some converted standard representation) and to present results of web queries in open and accessible data format representations. Metadata is used to document the existence of these datasets and the services available to access them. Reichman and Uhler (2006) state that “big science or mega science – NASA, NOAA, ESA, etc.” usually openly share their data and results in public repositories. However “small science” independent investigator driven research remains dominant in most scientific fields. Traditionally data from such studies have been extremely heterogeneous and un-standardised with few individuals making their datasets available through public repositories or even openly sharing them. In many “mashup”

creations using Google Maps (for example) non-GIS skilled developers have implemented methods called “screen scraping”. This involves extracting data from digital documents or maps in an non-automated manner or without the use of a formal conversion schema from the old model to the new model. Transformation errors commonly occur and consequently the accuracy of the newly created dataset may be difficult to verify. The data or information being scraped is usually neither documented nor structured for convenient parsing.

5.2 Closing Remarks

Some authors remain sceptical regarding the actual impact of web-GIS on the public. Kingston (2007) remarks that “there is still little evidence available as to how much the public are using such systems”. The author goes on to state that the vast amounts of money invested in e-Government a reluctance by local and national government to divulge data on access and usage leads one to suspect “that current usage of web-GIS for e-Government is relatively low compared with more traditional methods”. Mooney & Winstanley 2007c emphasis the importance of using web-log usage patterns taken from server machines running web-based geospatial services and analysing this information to assist in delivering better quality of service to the end-user and target areas of specific weakness. Combining spatial and non-spatial data presents unique challenges to data management and access over the web. Few web-based GIS environments have tools "out-of-the-box" for importing data and validating that data against a metadata profile as it is loaded into a database. Presently end-users must manually check if the coordinate system specified in the metadata is equivalent to the coordinate system in the corresponding dataset.

We feel that there will not be a sudden demise of desktop GIS in favour of web-based GIS in the foreseeable future. While the cartographical visualisations and interactivity of web-based GIS are continually improving desktop-GIS will remain a core component of GIS. This is due in no small part to the ability of desktop GIS to run large scale GIS queries involving very large datasets and complex spatial algorithms and database queries. As discussed earlier web-based GIS remains restricted in this regard due mainly to bandwidth considerations and network latency. Desktop GIS will become more Internet integrated. Desktop GIS users will no longer download several large spatial datasets directly to their local hard-disk drive or network drive in order to perform some GIS tasks. Instead these users will use WFS to download the precise subsets of the larger datasets or WMS to retrieve map layers. It is at this point that metadata has one of its most crucial roles to play in informing potential users that spatial data exists and providing information to allow these users to make judgement on the dataset's fitness for purpose and fitness for usage.

Web-based GIS allow organisations to expand their geospatial data holdings without any interruption to the end user or data consumer. Legacy datasets can be converted to the agreed spatial data representations and made available for access from the web-based GIS. With some intermediate work on updating style sheets, colour schemes, etc to accommodate these additional datasets these newly available data resources are quickly available to the organisation's stakeholders. Web-based GIS must be highly scalable – successful applications take advantage of networks with high bandwidth while working efficiently to avoid problems with slow networks and low bandwidth. The storage and representation of geospatial data is important in this regard. Very large environmental datasets should be stored in such a manner that queries can be more efficient – smaller data download size and smaller data transfer requirements for OGC services. The main goal at least for the foreseeable future will be “harmonisation” through interoperability in a service-based architecture rather than full-blown harmonisation of the underlying data models. Representation issues will be prevalent in both cases but much more manageable and achievable by using harmonisation through interoperability services.

References

Breslin,P., Ormsby, T., Napoleon,E., and Frunzi, N. **Getting to Know ArcView** ESRI Press. 1999

Bulterman, D.C.A. **Is it time for a moratorium on metadata?** IEEE Multimedia Volume 11, Issue 4, Page(s): 10 - 17 Oct.-Dec. 2004

ESRI The ArcIMS Internet Map Server, <http://www.esri.com/software/arcgis/arcims/index.html> Last Accessed March 2008.

GeoServer, A JAVA Driven GIS Data Server <http://geoserver.org/display/GEOS/GeoServer+Home> Last Accessed March 2008

Han, J. and Kamber, M. **Data Mining: Concepts and Techniques**, 2nd ed. The Morgan Kaufmann Series in Data Management Systems, March 2006. ISBN 1-55860-901-6

Kingston, R. (2007) "Public participation in local policy decision-making: the role of web-based mapping". *The Cartographic Journal*, 44(2), 138-144.

Leach, P., Strong, D, and Wood, G. "Canadian National Consultation on Access to Scientific Research Data (NCASRD)". Web-based Report: Available at http://ncasrd-cnadr.scitech.gc.ca/tdfdesignreport_e.shtml (Last Accessed Sept 2007)

MapGuide, The MapGuide Open Source Platform, <http://mapguide.osgeo.org/> Last Accessed March 2008

MapServer, MapServer Open Source Development Environment for web-based GIS, University of Minnesota, USA. <http://mapserver.gis.umn.edu/> Last Accessed March 2008

Miller, H.J., and Wentz, E.A. "Representation and Spatial Analysis in Geographic Information Systems" *Annals of the Association of American Geographers*, 93(3), 2003, pp. 574–594

Mooney, P. and Winstanley, A.C: **Applying Recommendation Algorithms to Collections of Geospatial Metadata**. *The 9th International Conference on GeoComputation 2007*, NUI Maynooth, Ireland. September 2007a

Mooney, P. and Winstanley, A.C. **Improving Environmental Research Data Management**. *Enviro Info 2007*, Warsaw Poland. September 2007b

Mooney, P. and Winstanley, A.C: **Evaluating Interfaces to Publicly Available Environmental Information: Human Computer Interaction 2007: Human Interface, Part I**, HCII 2007, LNCS 4557, Beijing, China. July 2007c

Nebert, D. ed., "Developing Spatial Data Infrastructures", The SDI Cookbook, 2004; <http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf>.

NSB, "Long Lived Digital Data Collections: Enabling Research and Education in the 21st Century". *Report of the National Science Board. The National Science Foundation*, 4201 Wilson Boulevard, Arlington, Virginia 22230, USA. May 2005. Available Online as PDF <http://www.nsf.gov/pubs/2005/nsb0540/> (Last Accessed September 2007)

Pillmann, W., Geiger, W., and Voigt, K. "Survey of Environmental Informatics In Europe". *Journal of Environmental Modelling and Software*. (21) 1519-1527. 2006.

PostGIS, Spatial Database Extension for PostgreSQL Databases. <http://postgis.refractory.net/> Last Accessed March 2008

Sharl, A. "Towards the Geospatial Web: media Platforms for Managing Geotagged Knowledge Respositories" *The Geospatial Web – How Geo-Browsers, Social Software and the Web 2.0 are Shaping the Network Society*. Eds. A. Scharl, K. Tochtermann. London: Springer. 3-14. 2007

Stachowicz, S. "Geographical Data Sharing – Advantages of Web Based Technology to Local Government". *Proceedings of 10th EC GI & GIS Workshop ESDI State of the Art*, Warsaw, Poland. 23rd-25th 2004.

Voigt, J. "Guidance Document on Implementing GIS Elements of the Water Framework Directive". *European Commission EC&GIS WFD Working Group*. Available Online at <http://www.ec-gis.org/docs/F2305/GIS-GD.PDF> (PDF Document – Last Accessed Sept 2007)

Waller, M. and Sharpe, R. "Mind the Gap: Assessing Digital Preservation Needs in the UK" Report

Prepared for Digital Preservation Coalition. Available Online in PDF format at <http://www.dpconline.org> (Last Accessed September 2007)

Wilson, J. R. (2002). "**Data and Information Management Strategy and Plan of the Global Ocean Observing System (GOOS)**". *Intergovernmental Oceanographic Commission*, Paris, France, GOOS Report No. 13 (IOC/INF-1168) edition.