Introduction to Web-based GIS

by Tara Athan
tara_athan@alt2is.com
1-707-485-1198

1. Overview

Web-based GIS differs from traditional GIS or paper maps in several important ways that may at times be advantageous or disadvantageous.

- User Community. The internet user community has greater variability in equipment, training and purpose than the traditional users of paper maps or GIS, but many more users can be reached at low cost through the internet.
- Viewing platform. The computer screen is the viewing platform in web-based GIS. In a traditional GIS shop, large format printers are available for hard-copy output, allowing a larger extent to be viewed at high resolution. Web-based GIS attempts to overcome this disadvantage by supplying zoom and pan controls so the user may view a larger extent at low resolution and zoom in for high-resolution details. Export capabilities of low-cost applications are often limited to printing the current view at screen resolution, but some open source applications are developing high-resolution, large-format export and print capabilities.
- Online Data Sources. Internet geographic data services allow access to large volumes of data without having to maintain this data on one's own webserver. Some of these sources are catalogued at (Anonymous, 2007). There is variation in the reliability of these internet servers, however. Non-profit organizations with limited resources have few options and must account for this uncertainty in planning, allowing, for example, extra lead time between pre-survey planning and a scheduled field survey.

To compare the capabilities of web-based GIS software, I will consider a set of representative services:

1. maps as illustrations for online education;
2. community reporting of locations having a defined set of attributes, and triggering notification of the webmaster;
3. documentation of work accomplished through the web presentation of maps;
4. online collection and retrieval of data and effort using a central database stored on a web server.

Three-dimensional maps can be useful in attracting a certain audience (i.e Google Earth users), but in general 2D maps are sufficient for these purposes. Temporal changes may be important; if that is the case, this information can be captured in a series of layers or time-stamp field.

<u>Users</u>

The users of the web-based GIS services listed above vary considerably. I have identified the following end-user groups and linked them to the service categories:

A. general public having basic internet skills (1, 2)
B. students at various levels, from primary to tertiary, with basic to advanced internet skills (1, 2)
C. managers and professionals who will use the system occasionally, with basic internet skills (1,2,4)
D. professionals who will use the system often, with advanced internet skills and basic GIS skills (1-4)

In addition to the needs of these broad categories of users, there are accessibility needs that arise for

certain users such as the visually-impaired and color-blind.

Security is an important consideration when implementing a system relying on community input. The user community should be defined in advance and an appropriate user-authentication approach developed. Backup and reversion is an essential need in order to recover from the inevitable vandalism as well as accidental loss of data.

Data

Base maps, such as administrative boundaries, elevation, watersheds, rivers and lakes, roads, municipalities and satellite imagery, may be required to provide context for the data to be presented. Geographical maps and data layers for these are available for web-based GIS applications from a number of online sources. Maps are defined in this context as georeferenced images with metadata. The Web Map Service (WMS) (OGC, 2007d) standard developed by the Open Geospatial Consortium (OGC) (OGC, 2007e) requires metadata specifying the coordinate system/projection, extent and name. This information allows the image to be included as a raster layer in a map composition with a legend. Geographical data, as opposed to maps, is delivered on the web either in the traditional GIS data models (vector, raster, coverage), tag-structured data in a markup language such as GML (OGC, 2004) or KML (OGC, 2007a), or as geotagged objects (Daviel, 2007) such as photos, feeds or blog entries. Some services, such as Google Maps, provide basemaps that are subject to use restrictions (Google, 2007), including a limit on the number of requests and public access without charge to the service. Care must be taken to give proper attribution and avoid the violation of use restrictions associated with online geographic data.

Some organizations may wish to make their own data accessible to all who can benefit from it. The traditional approach is to publish the data files in an index, allowing users to download and view in the application of their choice. This approach leaves conversion between formats and reprojection to the user, tasks that are easy enough for an experienced GIS user but beyond the expertise of the general public. A better approach is to build the map on the web server and deliver it to the browser, allowing the user to manipulate the view interactively. More advanced solutions are Web Feature Service (WFS) (OGC, 2007c) and Web Coverage Service (WCS) (OGC, 2007b), jointly called W*S, which store the data once while allowing a user to download in the format, coordinate system and projection of their choice. These services can be consumed by mapping applications, providing access to users without GIS, while also allowing download by experienced GIS users.

An alternative approach that is gaining in popularity is to publish geographic data as KML, which can be downloaded and opened by Google Earth. Other mapping applications are starting to support this format as well. The drawback is that KML is considered a format for 3D viewing data, not a format for geographic data (Google Earth Staff, 2007). Therefore it does not have the capacity for metadata that is needed for proper documentation of geographic data. An advantage to publishing KML data, however, is that it is Search-Engine-Friendly, while a W*S is not. A hybrid strategy is to publish in a KML file that includes metadata with a URL to the W*S.

Special concerns arise when publishing data on the web, including security and protection of privacy. A web-based GIS application may implement user-authentication to restrict access to data. To protect privacy, data may be generalized so that it is not possible to link attributes with particular parcels.

Software

I conducted internet and phone research, and attended the Free and Open Source Software for Geography 2007 Conference (Mitchell, 2007) in Victoria, Canada in order to learn more about software pertaining to web-based GIS. Before the conference, I had identified 15 relevant software products and at the conference I learned of 45 more. The open source software projects in particular tend to have modular components that are compatible with many other components. It would be too time consuming to evaluate all of these products in detail, so I have grouped them into categories with similar functions, as shown in Table 1. For each category, I have selected a type references- a relatively stable open source, free or cheap software product that best typifies the category. I then list all software products and services, that I am aware of, that have this capability, grouped into commercial, open source or services. Note that some products have a greater range of capabilities than others and may appear in more than one category.

Hardware

The main hardware decision in implementing web-based GIS services is whether to purchase and maintain a server or to subscribe to web hosting services. Most GIS server software support Linux and Windows platforms. If web hosting services are selected, a number of options are available:

- a shared hosting environment, like Bluehost (BlueHost, 2007);
- a shared open source GIS hosting environment, like MicroResources (Sherman, 2007);
- a dedicated open source GIS server, like HostGIS (HostGIS, 2007);
- a dedicated ArcIMS server, like Metropolis (Metropolis, 2007).

Aim and Scope of Report

In the following section, I identify several configuration options of software, hardware and data that match one or more of the services of interest listed above. These options will be compared with regard to capabilities, costs and risks, as well as the skills needed to implement the option.

2. The Feasibility Study

Option Details

Option I. Interactive web map presentation using OpenLayers or similar software, for service 1 and 3.

- The user sees a map window embedded within explanatory text. A familiar GUI allows interactive viewing including zoom, pan, layer selection through checkboxes in the legend and identify functionality through mouse events. Also provided is a link to download KML for offline viewing in a 3D Mapping Desktop Client, like Google Earth.
- Vector data is published on the web in two formats: KML and a format supported by the mapping software. This data appears in the map as an overlay on the basemaps obtained by WMS. Metadata is included in the data files to the extent allowed by the format standards, which is limited at present. A URL should be provided in the metadata leading to a detailed documentation of the data, or to WFS, if available.
- The content manager has responsibility for keeping the published data up-to-date and maintaining consistency among the published formats. The data files are edited offline, using desktop GIS and other applications, then uploaded using an FTP client or a file manager, like Cpanel, provided by the web hosting service.

- Administration of this web application requires knowledge of xhtml and simple Javascript, skills generally held by web designers and developers. Documentation and support through user forums is widely available for these languages, and numerous illustrative examples exist.
- Choices: there are many closed and open source applications capable of producing this option, as shown in Table 1. Characteristics to consider in the selection include: requirements for user-installation of plugins or components, cross-browser compatibility, compatibility of the license with the intended use, scalability, reliability of basemap source, quality of the user interface, quality of the cartography, maturity and stability of the application, and activity of the support and development communities.

Option II. Collaborative web mapping with e-mail reporting using OpenLayers or similar software, for service 2.

- The user sees two panels side-by-side, one containing an e-mail submission form, the other showing a 2D map as in Option 1. The map shows locations for the region. Attributes are selected from a form, which may include checkboxes, drop-down lists and text input. Location attributes are indicated by symbology; a legend provides clarification. More detailed information is accessible for each feature in a popup through a mouseover event. The two panels are linked so that when the user places markers on the map with a mouseclick, those geographic locations are added to the e-mail message. The user can also include text comments for each marker and in the e-mail body, can select a privacy option for the data and has an option to include contact information. Help links are available, including information and images of the species, detailed instructions on using the form and background information on the purpose of the data collection, and a submit button sends the e-mail.
- The e-mail recipient has responsibility for verifying the reported data if appropriate. If this person is not the content manager, then the report should be communicated to the content manager.
- The content manager has responsibility for updating the dataset to include the verified reports. In the simplest implementation, the dataset is edited offline and uploaded to the website as in Option II. A more advanced implementation would maintain the data in a geodatabase as described in Option III.
- E-mail functionality should be implemented server-side to allow usage from public computers at libraries and schools. Therefore, the developer must have knowledge of server-side processing in addition to web design skills as described in Option I.
- Choices: ideally, the software is the same application implemented and tested in Option I, as maintaining more than one application would require more training of staff and increase the potential for errors. The scripting language for the email handling must be selected: options include PHP, Perl, Javascript.

Option III. Web interface to PostGIS or other spatially-aware database, for service 4.

- The user logs in to the website to access a GUI for data entry, editing and query. The GUI is tailored to the dataset according to the schema. The registered user has limited permissions to change the data, and changes are tracked so that reversion is possible. The results of queries may be tabulated or shown visually in a map.
- The database administrator has full access to the database, and is responsible for maintaining user accounts, security, backups, modifying the schema if necessary, maintaining support tables as well as maintaining the application or scripts defining the GUI.
- Choices: the developer must select the web hosting service, the database software and the GUI

approach. Shared hosting services may support MySQL or PostGRES databases, but do not typically support the spatial extensions that allow the database to be used for geographic data. However, these databases without spatial extension would be suitable if only point data is used. If more complex data models (polygons, rasters) are employed, either a GIS hosting subscription would be obtained or a web server must be purchased and maintained. The cost of services and equipment, availability and expertise of staff and the organization's overall need for web service will inform this decision.

Implementation

Option I. Some organisations have a website and staff with sufficient skills to implement this option without assistance from a consultant, while others have no website or related expertise at present. Considerations which require a more professional implementation include making the maps available in older browsers that are not AJAX compatible or with javascript disabled.
Option II. Most small organizations are unlikely to have staff with sufficient expertise to develop this application, and so would require the services of a consultant to implement this option.
Option III. Extensive knowledge of databases and programming is required to implement this option.

Costs

I obtained costs of some of the commercial software and services from websites or quotes. A single ArcIMS license costs in the range of US$10,000 (ESRI staff, 2007), well out of the price range of a typical non-profit. An ArcIMS hosting service costs US$2000 per month (Metropolis, 2007), also not feasible. A shared GIS hosting service may be obtained for around US$20-30 per month (Sherman, 2007) while a dedicated server may be obtained for around US$100 per month (HostGIS, 2007). This is within the reach of most WMAs although this is more expensive than a shared host website, at US$10 per month (BlueHost, 2007). However, the extra cost seems well worth avoiding the installation of applications such as Map Server and GeoServer. Open source software is free to obtain, but can be challenging to install. I attempted for about 2 weeks to install Map Server on my shared host website, and although I obtained much advice from the developer's forum, in the end I was unsuccessful. I then acquired a GIS hosting service and was immediately able to create maps and data sources. The simplest options (1, 3) may be implemented on a shared hosting service as they do not require complex installations on the server.

3. Figures and Tables

Table 1. Categorization of web-mapping applications and services.

Category	Type Reference	Closed Source	Open Source	Service
Spatially-aware databases: store geographic features in a database, allowing queries by attribute and/or location.	PostGIS	ArcGIS Server, CubeSTOR, Oracle Spatial	MySQL, PostGIS	
Geographic data (WFS, WCS) servers: provide geographical data, vector and raster, to mapping clients through requests that specify the layer, data format, coordinate system/projection and extent desired.	GeoServer	ArcGIS Server, CubeSERV, ERMapper's Image Web Server	FeatureServer, MapGuide OpenSource Server, MapServer	
Web Map Service (WMS) servers: build maps for mapping clients through requests that specify the layers, coordinate system/projection, styles and image format desired.	Mapserver	ArcGIS Server, AspMap, CubeSERV, ERMapper's Image Integration Framework, Exposure Spatial/Image Server, GeoMedia's WebMap, MapXtreme, WebMap Composer	MapGuide OpenSource Server, MapServer, SharpMap	Internet Map Server, MapHub, Maporama
3D Web Mapping Engines: build 3D maps from geographical data in raster and vector form.	Virtual Terrain	ArcGIS Server, Exposure 3D	Virtual Terrain	

Web Mapping (W*S) Rich Internet Application (RIA) Client: display maps that combine online data from map or geographic data servers (W*S) with local data (on the same server as the webpage) without requiring installation of software on the client, minimal installation on the server and most of the processing performed on the client.	OpenLayers, Google Maps		ka-Map, MapGuide OpenSource Viewer (AJAX), msCross, OpenLayers, p.mapper, WorldKit	GoogleMaps, ArcWeb Services, MapQuest, MapTech, Mapufacture, MultiMap, Platial, Virtual Earth, Yahoo! Maps
Web Mapping (W*S) Client: perform the same task as the web mapping RIAs, described above, but require installation on the server and are thus difficult to implement in a shared hosting environment.	MapBuilder	Oracle MapViewer, pointMapper, TNTMapBuilder	Map Bender, MapBuilder, MapGuide OpenSource Viewer (DWF), Mapnik	
3D Web Mapping Desktop Client: display 3D maps with interactive controls.	Google Earth	ArcGIS Server, Google Earth (free version available)	MapGuide OpenSource Viewer (DWF), OSSIMplanet, WorldWind	
Web Geographic Processing Service (WPS) Server: perform geographic analysis, such as creation and editing of features, buffer and overlay, as a web application.	CartoWeb4	ArcGIS Server, AspMap, GeoMedia's WebMap	CartoWeb4, MapGuide OpenSource Server, PyWPS	

Desktop GIS: installed on the desktop, can also connect to W*S servers on the web, are an important tool for the webmap developer.	MapWindow, GRASS	ArcGIS Desktop, IDRISI, MapGuide Pro	GRASS, MapWindow, OpenEV, OpenJUMP, Quantum GIS, uDIG	
Spatially-aware Content Management Systems: CMS allowing content to be geotagged and displayed in maps.	Drupal		Drupal, Plone/PrimaGIS, Joomla/GMap	
Web GIS Development Tools: applications useful to the web map developer in creating map scripts and preparing data.	various	geoXtract (a free version is available)	Chameleon, EXIF, Mapnik, MapServer Workbench, MapStorer, Quantum GIS, shp2img, shp2KML	
Geographic Libraries, Installers, Utilities, etc.: programs and services installed on or called by the server in support of web map applications.	various	Geometry's Spatial Console, Map Vault	FWTools, GDAL/OGR, Geo::coder::US, GeoPress, GeoTag, GeoTools, GPSBabel, HostGIS, PROJ.4, Mapstraction, OSSIM, 52° North, FeatureServer	A2B, geocoder.us, GeoNames, GPSVisualizer

4. References

Anonymous. (2007) *WMS sites* [Online]. Available from World Wide Web: <http://wms-sites.com/> [Accessed Sep 30 2007].

BlueHost. (2007) *Web hosting provider* [Online]. BlueHost.Com. Available from World Wide Web: <http://www.bluehost.com/> [Accessed Sep 30 2007].

Daviel, A. (2007) *Geotags for HTML resource discovery* [Online]. Geosearch. Available from World Wide Web: <http://geotags.com/geo/> [Accessed Sep 30 2007].

ESRI staff (2007) *Personal communication*. ESRI.

Google. (2007) *Google maps terms and conditions* [Online]. Google. Available from World Wide Web: <http://maps.google.com/help/terms_maps.html> [Accessed Sep 30 2007].

Google Earth Staff (2007) *Personal communication*. FOSS4G2007, Victoria, CA: Google Earth.

HostGIS. (2007) *Dedicated servers with systems administration* [Online]. HostGIS. Available from World Wide Web: <http://www.hostgis.com/> [Accessed Sep 30 2007].

Metropolis. (2007) *ArcIMS hosting* [Online]. Metropolis New Media, Inc. Available from World Wide Web: <http://www.metropolisnewmedia.com/services/gishosting.html> [Accessed Sep 30 2007].

Mitchell, T. (2007) *Free and open source software for geospatial 2007* [Online]. OGC. Available from World Wide Web: <http://www.foss4g2007.org/> [Accessed Sep 30 2007].

OGC. (2004) *GML - the geography markup language* [Online]. Open Geospatial Consortium. Available from World Wide Web: <http://www.opengis.net/gml/> [Accessed Sep 30 2007].

OGC. (2007a) *KML overview* [Online]. Google. Available from World Wide Web: <http://code.google.com/apis/kml/documentation/> [Accessed Sep 30 2007].

OGC. (2007b) *Opengis® web coverage service (wcs) implementation specification* [Online]. Open Geospatial Consortium. Available from World Wide Web: <http://www.opengeospatial.org/standards/wcs> [Accessed Sep 30 2007].

OGC. (2007c) *Opengis® web feature service (WFS) implementation specification* [Online]. Open Geospatial Consortium. Available from World Wide Web: <http://www.opengeospatial.org/standards/wfs> [Accessed Sep 30 2007].

OGC. (2007d) *Opengis® web map service (WMS) implementation specification* [Online]. Open Geospatial Consortium. Available from World Wide Web: <http://www.opengeospatial.org/standards/wms> [Accessed Sep 30 2007].

OGC. (2007e) *Welcome to the OGC website* [Online]. Open Geospatial Consortium. Available from World Wide Web: <http://www.opengeospatial.org/> [Accessed Sep 30 2007].

Sherman, G. (2007) *Micro resources* [Online]. Micro Resources. Available from World Wide Web: <http://mrcc.com/> [Accessed Sep 30 2007].

