Francine Berish

Geolibraries, a review: the challenge of maintaining an organizational standard alongside the promotion of interoperability and access

Abstract: In the early nineties, academic debates surrounding geolibraries argued that the increased availability of data digitally via geo-spatial infrastructures would eliminate the role of the map librarian. Today map libraries and geolibraries not only co-exist, they are inter-connected. While several spatial data infrastructures have been developed and maintained since the mid-nineties, the amount of spatial data continues to increase via a number of domain specific and collaborative sources afforded by a Web 2.0. Current literature should remove its focus from developing data infrastructures, and direct its focus to modeling ways in which existing structures can be given incentives to become interoperable.

Key words: geolibrary, geo-spatial data infrastructure (GSDI), collaboration, interoperability, place

Introduction & Historical antecedents:

For the purpose of this study it is important to note the distinction between geolibraries and map collections/libraries. The map library's development is rooted in the wide distribution of war theatre maps following World War II (Parry, 2005). For the most part this growth took place within departments of geography. However, military utilization of government data using maps imparted the idea that maps were a viable source of information outside the field of geography (Parry, 2005). While geography libraries and geolibraries have a close relationship, the geolibrary does not necessarily have a spatial location. A geolibrary is one filled with geo-referenced information and based upon the notion that information may have a geographic footprint (Goodchild, 1998). However, the idea of a spatial footprint is not limited to maps; any form of media that can be given a locational variable is considered a footprint. For example, photographs, videos and music may be georeferenced. A locational footprint is often the primary key for database searching (Goodchild, 1998). Queries might resemble, 'Can I learn more about the neighbourhood where Elvis Presley was born and raised?' and 'Where can I purchase his music?' While these questions may be outside the perceived role of a geolibrary, it is important

to note that over eighty percent (80%) of digi-map use is outside the discipline of geography (Parry, 2005).

Although the geolibrary and the map library are defined as separate institutions, today they are far from mutually exclusive. In 1993, development of the internet and the analogous introduction and development of spatial data infrastructures brought into question the role of the map librarian as well as the future state of paper collections. Little academic literature until the late nineties reflected the idea that the geolibrary could coexist with traditional map libraries. The map librarian of today is the bridge between library and geolibrary.

While data may be increasingly available over the internet via data infrastructures, this does not mean it is organized or accessible. Promoting access – even digitally – is not outside the traditional realm of libraries. The management of digital data which is accessible remotely is not outside the realm of the librarian, "Libraries have always been about metadata (cataloging), information management (collection), access, sharing and preservation" (Boxall, 2003:20). While new technologies have enabled new methods of data management and storage for spatial information, the increasing quantity of spatial infrastructures does not discount the map library. There is a perception that maps, like books, are in decline as a result of the 'paper paradigm' (Keller, 2001), alongside decreasing investment in libraries, a result of the increasing cost of storage (Hawkins, 1998). Instead, digital geolibraries should be thought of as "...a new form of expressing very old institutions. 'Old' should not be equated with 'bad'; tradition has its place in modernity," (Boxall, 2003:19). The geolibrary can also be understood as a 'place'; an outlet for Geoscientists to meet with GIS librarians and geographers as well as other disciplines. "Libraries respond to many compiled societal needs. They are used for research, teaching, self-learning and entertainment" (Boxall, 2003:21).

In 1994 the National Spatial Data Infrastructure (NSDI) was developed by the Federal Geospatial Data Committee (FGDS) with the intention to create a "framework to implicate standards and policies to data sharing (public and private), across disciplines and between institutions" (FGDC, 2005). NSDI's seven geographic framework data themes are those which are typically generated by government agencies such as orthoimagery, elevation, transportation, geodetic control, hydrography, governmental units, and cadastral information (FGDC, 2005). Framework data is often the baseline for research, so, if made accessible, this data could be widely used by the public and private sector alike. The availability of said data could potentially contribute to more functional democratic processes, enhancing public participation and improving the effects and effectiveness of investments in spatial data (Onsrud, 2004). The mission of the NSDI is"…to reduce duplication of effort among agencies, improve quality and reduce costs related to geographic information, to make geographic data more accessible to the public, to increase the benefits of using available data, and to establish key partnerships…to increase data availability" (FGDC, 2005).

The Alexandria Digital Library (ADL) was established in the late 1990s as a response to various real and perceived problems of traditional map libraries, particularly access and organization (Goodchild, 2004). The ADL was established with the goal of creating the first geolibrary that could be searched by geographic coordinates. For example, placename queries are converted into coordinates using a gazetteer service. To develop this system, researchers at the University of California Santa Barbara created an automated catalogue system and a collection that could be accessed remotely. This acted as leverage in the investment in collections, since their static collection was largely unavailable to those outside the university (Goodchild, 2004). Unlike the traditional map library, "Digital storage would resolve issues of

preservation and the management of physical media" (Goodchild, 2004). ADL works "throughout the world to establish remote, independent, yet federated ADL nodes in which local collections can be added and maintained easily and effectively" (Masi, 2004), (Figure 1).

Emergent debates:

The progress of sharing over the internet has brought on debates about data availability and the role of librarian. The primary debate is founded on the issue of cost – should the cost burden be on the user or the publisher of data? "Spatial data can be characterized as downloadable good and one suited to trading in an electronic market" (Keenan, 2006; 4). However, simultaneously there is a perception that, "On the web the new frontier, everything must be shared, and everything free." (Campbell, 2000: 491). It is important to note when comparing case studies that national as well as international law plays an important role in determining data accessibility. The United States law states,

> …protection under this title is not available for any work of the United States Government, but the United States Government is not precluded from receiving and holding copyrights transferred to it by assignment, bequest, or otherwise.
>
> (Boxall, 2005, p 646)

Conversely, outside of the U.S., government or Crown Copyright regulates the distribution of data almost universally. For example, in Canada, Crown Copyright prevents the wide distribution of government data. To some extent, it is put into place to enforce an overarching standard of data quality. However, Crown Copyright also protects the government from being liable if their data is used for purposes for which it was not intended. Boxall (2005) claims that alongside the restrictions of Crown Copyright, Canada is a large area with a relatively small population, therefore organizations have fewer resources and longer travel times. This inhibits their ability to participate in the geo-spatial data infrastructure (GSDI) community.

While certain European nations have quickly realized the potential of making their data available via the internet, for large collections the cost and time alone to digitize only the existing collections is a burden. In 1996, the National Land Survey in Finland was the first to launch a site giving browsing and downloading access to its digitized topographic map series in its entirety (Parry, 2005: 198). However, in the case of large data holdings such as the British Ordinance Survey, scanning the twentieth century maps alone would take twenty-five years, not to mention the preceding years in which maps were individually catalogued, making them difficult to identify (Campbell, 2000). Although time is a large factor in large-scale digitization projects, the cost variable cannot be ignored. According to Campbell (2000), it would cost approximately four times the amount of the cost of the new building to digitize the British library's collection. The obvious question is who is going to foot the bill? While to some degree government information has been paid for with tax dollars, the funds are often allocated to collection, processing and distribution. Undertaking digitization projects and/or directing government distributors to GSDI's would demand additional spending.

Debate along the cost-time continuum highlights three emergent streams of thought on the debate of data accessibility. While the development of GDSI's have been influenced by disciplines outside of geography, it is clear that map librarians and geographers (and NeoGeographers) have the largest presence in the academic debate regarding data access. According to a study by Collingworth (2005) to gauge the skill set of librarians in the United Kingdom, ten out of eleven respondents had GIS capabilities in their libraries and over half provided user training. The debate of data access is not a debate divided by discipline, instead only by a varying degree of optimism on existing paper collections, who will facilitate access,

who will develop and organize it, and the feasibility of developing a GSDI for the large amount of spatial data to come.

Those who argue in favour of maintaining paper collections over digitization projects for the most part stay away from GIScience by exclusively talking about static scanned maps; however, what about digital access to vectorized data? (Parry, 2005). This leads the reader to further questions about whether the map curator wants to deal in information or in artifacts, "…provide the user with someone else's selection and presentation of data, but with the data itself and with the means by which the users can make their own selection and presentation of this data" (Tathman, 1994).

Ramifications:

To a large degree the cost debate and the 'paper paradigm' have had a large impact on existing paper collections. The cost of digitization and/or maintenance of the existing collections, in addition to processing new acquisitions, is immense. This would be alongside the responsibilities afforded by maintaining and facilitating access to digital collections. While the debate about data access has manifested in a way that has preserved the importance of the map librarian, fervency has surrounded the anticipated cost of providing services physically as well as digitally. As a result, in approximately thirty years, "…collections have been reduced in size, squeezed into tighter spaces, allowed to atrophy, or have been eliminated. And inevitably most have been starved for funds" (Parry, 2005). Academic literature published in the late-nineties was optimistic about the continued utility of paper collections; however, it is potentially rooted in fear of its *impending* redundancy in the digital future.

Today the multi-disciplinary roots of GSDI are recognized and the multi-dimensional skill set required by individuals working in map libraries and geolibraries is acknowledged in the academic debate. Despite the disappearance and disintegration of map libraries over the past thirty years, to some extent it is not realized that in practice the integration of paper map holdings with digital spatial data actually strengthens a library's collection (Parry, 2005). While spatial data accessible online could *potentially* be the same information represented on paper maps, developments of internet technology have increased the usability and ease of application development, marking the transition to a Web 2.0 (Haklay & al, 2008). This transition allowed users to contribute and modify spatial data remotely, a representation of spatial data distinct from the paper map and its digitized version, both in its organization and storage. Web 2.0 technologies have afforded recognition that geography, or location, is an important means to index and access information on the internet (Haklay & al, 2008).Web 2.0 is characterized by a bi-directional exchange of data allowing the user to manipulate and interact with maps.

Early scholars were quick to fear the top-down impact of commercial industry control regarding data access. Campbell (2000) took particular care to note the potential of grass-roots efforts. Even individuals could have an impact on the collection, input and distribution of data in the digital world. Campbell (2005) gives the example of Roelof Oddens who, in his spare time, updates his website containing a directory of links to websites containing 10,500 thematic maps, with approximately twenty websites added per day. This is a significant accomplishment for an individual. Campbell (2000) also claims that early retirement and longevity will have a significant increase in the number of researchers, thus of bottom-up contributions to GSDI's. While Campbell's argument is inspiring, it is equally difficult to fathom to the mobilization of the retired community for the promotion of data access. To some extent Campbell's prophecy

was self-fulfilling though through a different mode – the public at large out represent retirees. An example of collaboration afforded by Web 2.0 is Open Street Map (OSM) in the U.K., which has been generated on a voluntary basis, and is "a free editable map of the whole world. It is made by people like you" (OSM). OSM allows users to view, edit and use geographical data in a collaborative way from anywhere on Earth. Despite strict Crown Copyright regulations in the U.K., in order to have copyright on their data OSM is produced by volunteers who reproduce the street grid using global positioning systems (GPS) or GPS enabled devices such as cell phones. Since it is not simply a digitized crown map, it is understood as produced by OSM. However, OSM was a triumph of public participation, regulated through numerous samples and careful monitoring by devotees. The project was made possible by the technological development in personal GPS units which ultimately increased the number of samples and the extent of public participation within the project–especially in the developed world.

Conclusion:

While the early academic debate speculated a reversal of roles in the preparation, organization, ordering and handling of data via the internet, librarians have safeguarded their role in data processing. To some extent, static map collections have decreased in size and resources. However, map librarians play a role in digital collections that are remotely accessible. Remote accessibility and an increase in users necessitate the role of a human intermediary. Today, to a librarian with a multi-disciplinary skill set in GIS, computer science, library cataloguing and database management, "There is no doubt that users of digital data, especially where the data are expressed as a map, often require the services of people who can help them acquire, download, view, print, or even interpret the data or the map" (Parry, 2005). Despite the future forecast

portrayed at the onset of GSDIs, the map library and the geolibrary exist simultaneously. Even though a geolibrary does not have a spatial location, according to Hawkins and Battin (1998) geolibraries are more than just storage facilities; "they are active learning places and long-standing contributors to the economic and social vibrancy of our communities". Since 2006 there has been little academic literature on the future dynamic of map libraries and geolibraries. Instead, much of the current literature has focused on the development of geolibraries by public and private institutions alike. While existing GSDIs like the NSDI and ADL (Figure 1) have continued to grow, an increase in data has also resulted in an increased number of databases, making the prospect of data standardization and quality control seem unlikely. Instead of concentrating on the development of GSDIs, current literature should now address the problem of interoperability between infrastructures on a large scale. The large quantity of data created by users and afforded by a Web 2.0—map mash-ups, crowdsourcing, mapping application programming interfaces (API), NeoGeography, geostack and tags—must also be addressed. Allen (2008) has demonstrated the possibility of developing standard procedures such that the collection of metadata from paper maps in a library collection may be integrated into the NSDI. While the subject of interoperability has been addressed at a small scale in case-specific examples, there must also be literature on incentives for data infrastructures to adopt said models. The impact of quickly changing technologies and increased data storage is difficult to forecast, therefore at this time, "The only certainty about the future is its uncertainty" (Boxall, 2003).

Appendix:

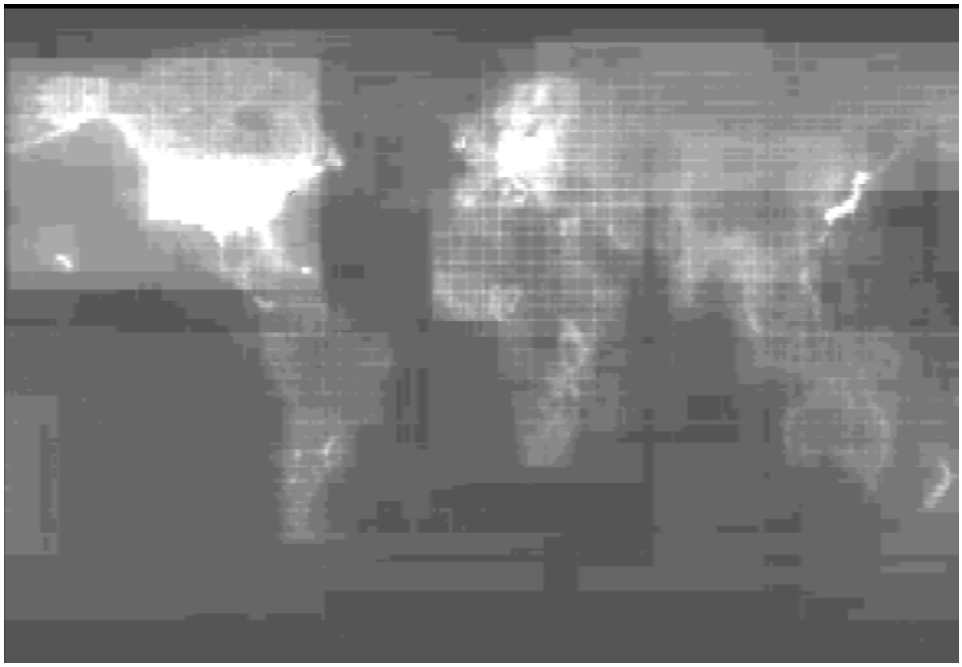

Figure 1. Collection Level Meta-Data of Alexandria Digital Library (Goodchild, 2002).

Reference List:

Allen, Marcy 2008: Bridging the Gap Between Spatial Data Infrastructures and the Map Libraries, *OCLC Systems & Services: International digital library perspectives*, 24(3):153-159.

Boxall, James 2003: Geolibraries: geographers, librarians and spatial collaboration, *The Canadian Geographer*, 47 (1): 18-27.

Boxall, J 2006: The nature of geospatial information and security, *Government Information Quarterly*, 22: 644-662.

Campbell, Tony 2000: Where are Map Libraries Heading? Some Route Maps for the Digital Future, *LIBER QUARTERLY*, 10: 489-498.

Collingworth, C. 2005: Implementation of geographic information systems (GIS) into library and archival systems, *Bulletin, Association of Canadian Map Libraries and Archives*, 122: 3–16.

Federal Geographic Data Committee (FGDC) 2005: The National Spatial Data Infrastructure. Online: http://www.fgdc.gov/nsdi/library/factsheets/documents/nsdi.pdf. Accessed on: 29/01/2010.

Goodchild, Michael 2004: The Alexandria Digital Library Project; review, assessment and prospects, *D-Lib Magazine*, 10 (5): 1-9.

Goodchild, Michael. 2002: Data Access and Data Sharing Online: http://www.csiss.org/aboutus/presentations/files/ Accessed on: 01/02/2010.

Haklay, M., Singleton A. and C. Parker (2008): Web Mapping 2.0: The Neogeography of the GeoWeb. In *Geography Compass*, 2 (6): 2011-2039.

Hawkins, & P. Battin (1998): The mirage of continuity: reconfiguring academic information resources for the 21st century, *Washington, DC: Council on Library and Information Resources and the Association ofAmerican Universities*, 129-163.

Keenan, Peter 2006: An Electronic Market for Spatial Data, *Proceedings of the Twelfth Americas Conference on Information Systems, Acapulco, Mexico August 04th-06th 2006*: 1-4.

Masi, C. 2004: Alexandria Digital Library. Online: http://www.alexandria.ucsb.edu/adl/about_adl.html#History Accessed on: 02/02/2010.

Onsrud, Harlan 2004: The Future of the Spatial Information Infrastructure, in A research agenda for geographic information science: 225-255.

Open Street Map (OSM): The Free Wiki World Map. Online: http://openstreetmap.org Accessed on: 01/20/2010.

Parry, R.B 2005: Moving On. Where Should the University Map Collection be going? *The Cartographic Journal*, 42 (2): 197-201.

Tatham, Andrew (1995): Can the map curator adapt*? The Liber quarterly*, 5, (3): 330-336.