A Matter of Degrees:

Improving access to geospatial information in libraries through geographic searching

Sarah Simpkin
LIS 9410 Independent Study
Supervised by Victoria Rubin
5 August 2011

Table of Contents

Introduction

The intersection of information retrieval and Geographic Information Systems (GIS) in recent decades, particularly as a result of the integration of coordinate information within catalogue records, has produced a wealth of innovative services built to query records using geography as the primary search mechanism. This paper first presents an overview of the literature published on this topic in recent decades. It then presents the results of a study intended to explore those catalogue record fields that can encode geographic information about maps and their content. This paper will shed some light onto the current usage of these fields, as well as their utility in the future as access points for geospatial information.

Literature Review

This literature review traces a history of developments in spatial searching by first exploring the challenges and limitations inherent in cataloguing maps and other spatial information. Some of the advances made possible through the inclusion of coordinate information in bibliographic records are then discussed. A number of ongoing challenges pertaining to geographic searching are explored. Finally, the literature review concludes with a list of predictions about the future of spatial searching. A guide to some of the terminology employed in this paper is included as Appendix A.

Cartographic cataloguing challenges

Much of the criticism surrounding today's map cataloguing practices stems from the development of MAchine Readable Cataloguing (MARC) by the Library of Congress in the late 1960s. In an effort to ensure compatibility between the records produced by each division of the Library, it was determined that data ought to be entered into MARC records in accordance with the Anglo-American Cataloguing Rules (AACR), a standardized set of cataloguing rules published in 1967 (Webster 1982, 61). The AACR requirement proved challenging for map cataloguers, who argued that the rules (based largely on monograph cataloguing guidelines) failed to capture the unique idiosyncrasies of cartographic materials[1] (Webster 1982, 62). Among these, some of the strongest criticisms raised by map librarians concerned the AACR stipulation that catalogue main entries should reflect the material's author rather than the area covered by the map itself (Stibbe 1976, 41). Indeed, this practice has been maintained in the current, second edition of AACR (Moore and Hall 2001, 5). As digital records have replaced card catalogues, however, concerns over the choice of main entry have diminished in recent years (Larsgaard 1998, 168) and MARC has become a widely-accepted standard across map libraries (Larsgaard 1998, 185). As of 1999,

---

[1] By the time the second edition of the rules (AACR2) was published in 1978, however, the Library of Congress Geography & Maps division had proposed over 50 changes to the original rules (Larsgaard 1998, 168).

56.6% of research and academic libraries in North America reported at least 70% MARC records in their map catalogues (Davis and Chervinko 1999, 23).

The widespread adoption of AACR (now in its second edition – AACR2) and MARC by map libraries has exposed additional cataloguing challenges beyond the main entry question. These include:

1. The lack of support in AACR2 for handling map series, preventing the linking of individual sheet maps within a series to the series record as a whole (Larsgaard 1998, 190; Rockwell 1999, 53; Webster 1982, 62).
2. The lack of direction provided by AACR2 concerning the cataloguing of multiple maps printed on one sheet (Larsgaard 1998, 196; Rockwell 1999, 52). Should each map be catalogued separately, or should a single record be produced only for the most prominent map?
3. The emphasis on dates of publication in AACR2 rather than dates of information, which are typically more relevant to researchers (Larsgaard 1998, 189). Relatedly, the lack of clarity provided by some map publishers as to the currency of their data may prove challenging to cataloguers (Rockwell 1999, 50).
4. The dearth of standards in place for describing coordinate information in catalogue records and the inconsistent use of fields designated for this purpose. For example, the optional MARC field for coded cartographic mathematical data (MARC 034) allows cataloguers to enter the bounding coordinates for the extent of each map as either HDDDMMSS (Hemisphere-Degrees-Minutes-Seconds) or as decimal degrees (Gonzalez 2007, 6; Moore and Hall 2001).
5. The failure of Minimal Level Cataloguing (MLC) as defined in AACR2 (Rule 1.0D1) to capture essential components of map description. These include a failure to mandate the inclusion of map dimensions (MARC 300 $c), coordinate information (MARC 034, 255 $c), and multiple subject headings, among others. These missing fields may pose significant barriers to the access of geospatial information (Ercegovac 1998).

Spatial searching and the value of coordinate data

The value of providing access to cartographic materials and other geospatial data by taking advantage of the spatial information embedded within bibliographic records has been recognized for decades. Writing in 1967, Kate Donkin and Michael Goodchild described a coding system for map records implemented at McMaster University which enabled users to search for materials using place names. The nature of the codes meant that records for maps featuring nearby areas could be displayed in close proximity to one another in the catalogue (Donkin and Goodchild 1967, 41).

More recent work in the field of geographic searching has focused on the challenges inherent in offering “human-friendly” search techniques for maps and geospatial data. Since entering raw coordinate information into a search field is a non-intuitive process for most users (Yu 1999, 253), researchers have been focussing in recent years on creating more instinctive ways of retrieving maps from digital databases. These spatial queries may take the form of selecting the desired area on a graphical map representation or by specifying a particular place name or political jurisdiction in a text search (Gonzalez 2007, 9).

Place names employed as subject headings (MARC 6## $z) have excellent potential as access points for cartographic data but face a number of shortcomings in their current form. These include ambiguity[2], differing transliterations into Roman scripts[3], different names in different languages[4], name changes[5], anachronisms[6], and changing footprints over time[7] (Buckland et al. n.d., 3, Vestavik 2004, 2). Gazetteers—geographical indexes providing place names, place categories, and their associated coordinates—offer one solution. Using digital gazetteer services to tag cartographic records with coordinate information unlocks a host of geographic searching possibilities, including the ability to deduce relationships (for instance, adjacency, containment, and overlap; see Figure 1) between objects within the collection (Gonzalez 2007; Goodchild 2004). Topological relationships inherent in gazetteers may also be exploited for indexing and searching purposes[8] (Riekert 2002, 587).

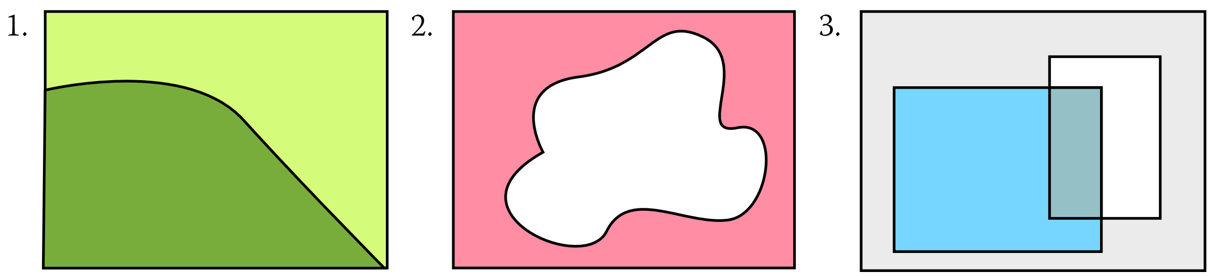

Figure 1. *Geographic searching may enable certain topological relationships between objects to be deduced. Here, adjacency (1), containment (2), and overlap (3) are illustrated.*

Additionally, recent changes to the coded cartographic mathematical data field (MARC 034), introduced in 2006, allow start and end dates ($x and $y, respectively) to be recorded, providing support for changing boundaries over time (Library of Congress, 2006). It is hoped by the MARC

[2] *Ontario,* a Canadian province, or *Ontario*, a city in California?

[3] *Peking* or *Beijing*? *Calcutta* or *Kolkata*?

[4] *Deutschland, Germany,* or *Allemagne*? *Nova Scotia* or *Nouvelle-Écosse*?

[5] *St. Petersburg* to *Leningrad*, then back to *St. Petersburg.*

[6] *Austria-Hungary* became a much smaller *Austria.*

[7] Cities may expand over time, amalgamating with surrounding communities.

[8] For instance, boundary information from one jurisdiction may be used to infer whether a particular city falls within its borders.

Advisory Committee that these amendments to field 034 (among others) will lay the groundwork for coordinate-based retrieval of all records containing geographic terms (Library of Congress, 2006). This expansion of geographic searching to non-cartographic materials is echoed in Michael Goodchild's use of the term geolibrary, a library whose search mechanism for all materials (photographs, news stories, music, etc.) is based primarily on geographic location (2004). Indeed, there are numerous MARC fields (non-exclusive to map records) that might be mined successfully for geographic clues. A study conducted by Vivien Petras of the University of California library system revealed that almost half of the 5 million MARC records examined contained geographic codes (MARC 043 $a), with geographic subdivision added entries (MARC 650 $z) and geographic subject added entries (MARC 651 $a) employed in 35% and 18% of the records respectively (2004, 1). There exists a great potential for the expansion of geographic searching to include a wide range of materials not traditionally ascribed geographic coordinates.

One early pioneer of the geolibrary concept is the Alexandria Digital Library Project (ADL) based out of the University of California, Santa Barbara. Founded in 1994, the ADL offers a graphical map interface for interacting with the collection of spatial data and airphotos, whose records are all MARC-compliant (Goodchild 2004). The search process is further supplemented by an extensive gazetteer service, enabling place name-based retrieval (Goodchild 2004). User studies conducted in the early years of the project provide some insight into the service's target user groups and their experiences with multiple iterations of the system (Hill et al. 2000, 252). It was determined that future iterations of the interface ought to provide a greater number of search functions (notably, the ability to select non-contiguous search areas and the ability to sort results by type, date, etc.) while presenting a unified search area that integrates gazetteer-, catalogue-, and map-based searching (Hill et al. 2000, 257). It seems clear that users on the whole are demanding well-integrated search functionality alongside an intuitive, immersive map browsing experience.

Another relatively early cartographic information retrieval system, GeoMatch, was conceived in part to determine whether quantitative comparisons might be drawn between user-specified search areas and the retrieved map coverages (Yu 1999, 258). The results showed that it was possible to rank the search results by the degree to which the two areas overlapped (Yu 1999, 258). Indeed, the emergence of GIS concepts within information retrieval literature points to the growing popularity of geographic searching within the library community.

Ongoing challenges in geographic searching

Fundamental to the idea of geographic searching is its potential to combine two very different access models for retrieving information (Larson 1996, 83). Probabilistic searching, common in information retrieval, endeavors to pair users with the most relevant materials as assessed by the user in any given situation, while deterministic searching has much more rigid outcome—materials are relevant if they fulfill the conditions specified at the outset (Larson 1996, 83). Geographic searching accomplishes both—probabilistic queries may produce a list of data sets which contain information on a particular pair of coordinates, while deterministic queries may return a listing of all cities within a specified radius (Larson 1996, 83). The most versatile systems should therefore offer support for both access models.

Another major challenge in providing useful geographic searching services to users is the failure by most gazetteers to encode the spatial and semantic relationships between places . According to Vestavik (2004, 4), hierarchical relationships between objects[9] are very infrequently included in gazetteer services but remain a fundamental way in which humans organize spatial information . Since thesauri are traditionally employed in information retrieval to organize relationships between terms, their inclusion within geographic searching services (as a supplement to gazetteers) is a logical next step .

Next, the growing popularity of browser-based mapping services (Google Maps, Google Earth, Openstreetmap, and Bing Maps, among others) is changing the typical user profile for geospatial data resources. Today's "geo-hackers", web-savvy users who may not possess a formal background in GIS, are producing interactive maps online using data from a wide variety of sources, otherwise known as mashups (Morris 2006, 295). Unlike traditional geospatial data users, who typically require access to a large number of resources but whose audiences tend to be relatively small, these new developers are reaching large numbers of people with their online mapping projects (Morris 2006, 295). These users, who may initially be unaware of data quality standards, are increasingly seeking advice on the selection of data sources for their maps (Morris 2006, 295). Indeed, the public at large are increasingly being exposed to a plethora of mapping tools and services, including portable GPS units, online route planners, and address locators, among many others (Hill 2000, 257). As a result of this growth, it is imperative for libraries to anticipate the emerging needs of these non-expert users.

Similarly, users accustomed to online mapping services may have particular expectations when it comes to the online browsing experience for geospatial data. Google Maps has been a groundbreaking player in this respect. The browser-based service, which endeavors to provide an "interactive, exploration-oriented user modality" (Jones 2007, 9), encourages users to discover information spatially through a combination of text searches and a large interactive map which dominates the browser space. Even more immersive is the Google Earth application, in which a three-dimensional visualization of the earth becomes a vivid backdrop for georeferenced information shown in situ (Jones 2007, 11). This hands-on, visual approach to information discovery is often at odds with traditional online catalogues for maps and geospatial data, which tend toward text-based lists rather than explorable, map-centered displays. In this respect, services like the ADL, which feature a graphical map interface for exploring the extents of datasets, are a step in the right direction.

---

[9] For instance, Signal Hill is *within* St. John's, which is *within* the St. John's Metropolitan Area, *which* is *within* Newfoundland and Labrador, which is *within* Canada, etc.

Additionally, there are some more technical obstacles to overcome concerning the mechanics of retrieving spatial information. Notably, there exists a certain ambiguity concerning the coordinate information presented in catalogue records. Briefly, coordinates are often provided to represent the four corners of a typically rectangular map or dataset, a "bounding box" enclosing the data within (Caldwell 2005). However, it is often unclear to cataloguers whether these coordinates are representative of the entirety of the data set (the data extremities), or simply the study area within a larger spatial extent (Caldwell 2005). While this uncertainty is doubtless more common when dealing with digital data rather than maps whose full extents are clear on the page, this situation does warrant some attention (see Figure 2).

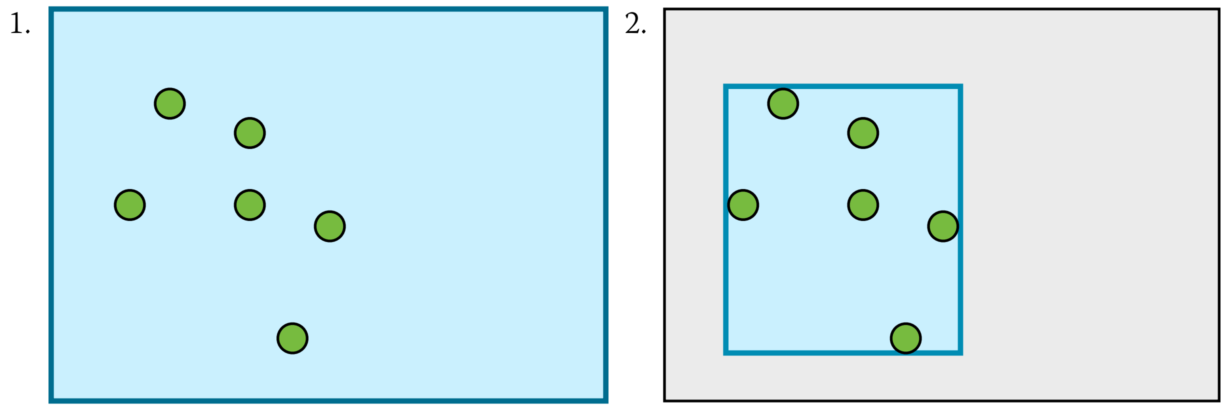

Figure 2. *Bounding box coordinates for a dataset may be reported for the entire study area (1) or strictly for the outermost extents of the data points contained within it (2). This ambiguity can introduce problems when creating graphical indexes of the dataset.*

Looking ahead

Based on the work described above, several predictions can be made about the future of geographical searching and the retrieval of spatial information.

1. Using coordinate information as a basis for geographical searching will enable map libraries to overcome some of the challenges associated with changing political and administrative divisions (Gonzalez 2007). Relatedly, recent additions to the 034 field in MARC will note start and end dates for spatial objects in an effort to overcome the problems associated with changing jurisdictions and boundaries (Library of Congress 2006). The absence of 034 as a required field in the minimal level cataloguing guidelines for cartographic materials is a step backward in this respect.
2. Geolibraries will enable the expansion of geographic searching to many kinds of materials, not just cartographic ones (Gonzalez 2007; Goodchild 2004).
3. Geographical searching is redefining the way we can query spatial databases. GIS concepts like proximity, containment, and adjacency are making their way into the information retrieval literature (Gonzalez 2007).

4. Geographers are not the only people looking for spatial information. The popularity of web mapping services like Google Maps and Google Earth has raised the profile and market for geospatial data (Morris 2006, 295). Libraries must prepare for the increasing geospatial needs of non-experts.
5. Relatedly, the increasing demand for digital geospatial data raises questions about the long-term management of these resources. Do libraries have a role to play in the preservation of digital datasets? (Morris 2006, 296). It is especially vital to consider that not all users are looking for the very latest data, particularly when performing time-series analyses and historical research.

Geographic searching is undeniably a growing area at the juncture between GIS, information retrieval, and computer science. Numerous advances made in recent decades, particularly web-based systems that integrate coordinate information found within bibliographic records, gazetteer data, and thesauri, are changing the ways in which users can visualize and access spatial data. In spite of these developments, however, geographic searching remains limited by the non-mandatory reporting of coordinate information in AACR2, questions about how bounding boxes represent data boundaries, and the difficulty of creating user-friendly, intuitive search interfaces that preserve the semantic relationships between objects.

Study Objectives

Building on the spatial searching literature surveyed above, this study was designed to evaluate how thoroughly spatial information is encoded in existing catalogue records. Which MARC fields containing geographic information are most commonly used by cataloguers? Do these fields provide enough information about the maps to infer which geographic areas are covered by their extents? Finally, how easily can these records be adapted to suit changing place names and anachronisms?

Methods

Selection criteria

To better understand how geographic information is encoded and updated within bibliographic records, a convenience sample of ten maps exemplifying some of the cataloguing challenges (most notably, name changes and anachronisms) listed by Buckland et al. (n.d.) was produced. Since our intention was to compare how these issues are addressed in different cataloguing practices, we retrieved the MARC record for each map from two Canadian academic library online catalogues: The University of Toronto and the TriUniversity Group (Universities of Guelph, Waterloo, and Wilfrid Laurier), hereafter known as University A and University B. In creating the dataset, precedence was granted to maps that were both available in the two library catalogues and those which featured places whose names or extents had changed since the maps were first published.

A data collection instrument was then designed to test for the presence or absence of cartographic and geographic coordinate fields (in MARC 034 and 255), as well as for the additional fields suggested by Petras (2004, see Appendix B) that might contain geographic "clues" about the maps themselves. Fields on the list which were present and contained information were coded with a "Y" for "yes". Empty spaces corresponded to the absence of cataloguing information for that particular field.

Procedure

Records were located using the online catalogues for each library and displayed using the "MARC view" feature on both websites. A screenshot of each record was then saved for later viewing. A side-by-side example of one map's record in both libraries is shown in Appendix C. Our data collection instrument consisted of a Microsoft Excel spreadsheet (see an excerpt in Appendix D). Maps were listed across the top of the page and the MARC fields of interest were enumerated in the first column. In this configuration, fields coded with "yes" were rapidly distinguishable from the rest.

Results

The first series of fields to be tested were those proposed by Petras (2004, 1-2) deemed to contain potential geographic evidence of the maps' coverage areas (and to some degree, their place of publication). Of these 17 fields, only three were used consistently (see Figure 3). These consisted of MARC 008 $15-$17 (Place of publication, production or execution), 052 (Library of Congress geographic classification), and 260 $a (Place of publication, distribution, etc.). Of these, fields 008 and 260 are considered mandatory according to the Association of Canadian Map Libraries and Archives' Core Level Cataloguing document for non-serial cartographic materials (ACMLA 2008). In descending order by use, field 651 $a (Subject added entry-geographic name) was completed in 85% of cases, 650 $z (Subject added entry - topical term - geographic subdivision) in 30% of cases, and 043 $a (Geographic area code) in just 5% of cases. No maps in the sample dataset contained any information in fields 033 $b-$c (Date/time and place of an event), 518 $a (Date/time and place of an event note), or 522 $a (Geographic coverage note). Additionally, there were no major differences in field use by institution.

		University A		University B		Combined	
Field	**Field Name**	**#**	**%**	**#**	**%**	**#**	**%**
008	*Control field 008*						
$15-17	Place of publication, production or execution	10	100	10	100	20	100
033	*Date/Time and Place of an event*						
$b	Geographic classification area code	0	0	0	0	0	0
$c	Geographic classification subarea code	0	0	0	0	0	0
043	*Geographic area code*						
$a	Geographic area code	1	10	0	0	1	5
$b	Local GAC code	0	0	0	0	0	0
$c	ISO code	0	0	0	0	0	0
044	*Country of publishing/producing entity code*						
$a	MARC country code	0	0	0	0	0	0
$b	Local subentity code	0	0	0	0	0	0
$c	ISO country code	0	0	0	0	0	0
052	*Geographic classification*	10	100	10	100	20	100
260	*Publication, distribution, etc. (Imprint)*						
$a	Place of publication, distribution, etc.	10	100	10	100	20	100
518	*Date/time and place of an event note*						
$a	Date/time and place of an event note	0	0	0	0	0	0
522	*Geographic coverage note*						
$a	Geographic coverage note	0	0	0	0	0	0
650	*Subject added entry – topical term*						
$c	Location of event	0	0	0	0	0	0
$z	Geographic subdivision	4	40	2	20	6	30
651	*Subject added entry – geographic name*						
$a	Geographic name	9	90	8	80	17	85
$z	Geographic subdivision	0	0	0	0	0	0

Figure 3. *This table shows the use of fields featuring geographic "clues" as per Petras (2004, 1-2). The number of times these fields were used to describe the ten-map sample are included for each university in the study along with their corresponding percentages.*

In the case of fields containing explicit coordinate information and other cartographic data, the two institutions differed in the extent to which they included certain fields (see Figure 4). Most notably, University A included MARC 034 $a-$b (Category of scale and constant ratio linear horizontal scale) in all ten records while University B provided them in only 30% of cases. Curiously, these fields are considered mandatory by ACMLA's core-level cataloguing guide (2008). Coded coordinate fields 034 $d-$g are considered optional by ACMLA (2008) but were however included in 40% of the records examined. Finally, cataloguers were relatively thorough when populating fields 255 $a-$c, reporting scale, projection, and coordinate information in 95%, 50%, and 40% of records respectively.

		University A		University B		Combined	
Field	**Field Name**	**#**	**%**	**#**	**%**	**#**	**%**
034	*Coded Cartographic Mathematical Data*						
$a	Category of scale	10	100	3	30	13	65
$b	Constant ratio linear horizontal scale	10	100	3	30	13	65
$c	Constant ratio linear vertical scale	0	0	0	0	0	0
$d	Coordinates - westernmost longitude	5	50	3	30	8	40
$e	Coordinates - easternmost longitude	5	50	3	30	8	40
$f	Coordinates - northernmost latitude	5	50	3	30	8	40
$g	Coordinates - southernmost latitude	5	50	3	30	8	40
255	*Cartographic mathematical data*						
$a	Statement of scale	10	100	9	90	19	95
$b	Statement of projection	5	50	5	50	10	50
$c	Statement of coordinates	5	50	3	30	8	40

Figure 4. *The use of fields featuring cartographic and coordinate information. Records were examined for the ten-map sample and displayed as counts and percentages.*

Discussion

Many of our findings in terms of geographic clues encoded within catalogue records are consistent with those reported by Petras (2004, 1-2). The complete adoption of MARC fields 008 (Place of publication, production, or execution) and 260 $a (Place of publication, distribution, etc.) as well as the absence of fields 033 $b-$c, 043 $b-$c, 044 $a-$c, 518 $a, and 544 $a are also reflected in her results. On the other hand, while our two university libraries employed the Subject added entry - Geographic name field (651 $a) in 85% of cases, only 18.14% of records in the University of California's catalogue seemed to do so (Petras 2004, 1-2). This may be attributable to the fact that our study focused solely on map cataloguing practices and not those for a wider breadth of materials. Interestingly, however, both studies revealed similar usage rates for field 650 $z (Subject added entry - topical term - geographic subdivision) of 30-35%. Should these geographic subdivisions ever be tagged with coordinate information, a host of spatial searching possibilities would be unlocked (Goodchild 2004; Library of Congress 2006). That almost one third of records in the study contained geographic subdivision information is therefore a very promising finding.

The presence of coded coordinate information (MARC 034 $d-$g) in 40% of map records is another positive step toward the creation of online map browsers and other tools to discover–and interact with–geospatial information within a collection. While these fields remain optional according to the ACMLA core level cataloguing guidelines, their presence should greatly facilitate the process of creating tools like online graphical indexes which take into account the spatial extent of each map. Nevertheless, cataloguers should be aware of two shortcomings concerning these fields. First, confusion may be introduced since coordinates can be entered in both decimal degree and hemisphere-degree-minute-second formats (Gonzalez 2007, 6; Moore and Hall 2001). Secondly, the bounding box problem could arise should a map's outermost coordinates not correspond to the extent of the data found within the map itself (Caldwell 2005). Ultimately, however, MARC 034 holds great promise as a way to perform geographic searches using existing maps in a collection.

Finally, our selection of maps featuring anachronistic place names and changing borders provided a great deal of insight into how records might benefit from more granular geographic added entries. One example in the study concerns a 1959 map of Otto Fiord Glacier on Baffin Island. At the time of publication, the entire island was part of Canada's Northwest Territories (NWT). However, these lands became part of the newly-designated Nunavut territory in 1999. At University A, MARC 651 $a contained an entry for the corrected "Baffin Island (Nunavut)", while University B's record reflects the glacier's former position within NWT. In this instance, the use of a more granular entry (Baffin Island) by University A likely enabled the record to be easily updated to reflect the boundary change. These simple changes to 651 may also have facilitated updates to both universities' records of a 1973 map of Zaire, now called the Democratic Republic of Congo. These updates were not a universal phenomenon, however. Users searching for historical maps of Thunder Bay, Ontario might not be aware that the 1964 map of Port Arthur covers part of the same area; Port Arthur and Fort William amalgamated to form Thunder Bay in 1970 but no evidence of the new name appears in any part of either university's record.

Conclusion

The geographic "clues" found within catalogue records for cartographic materials have the potential to transform how users interact with spatial data online. Many fields contain references to the extent covered by each map that exceed the information contained within main entries and map titles. Online tools that can take advantage of these fields–particularly those containing coordinate information and geographic added entries–are becoming the new standard for discovering geospatial data. At a time when expert and non-expert users are increasingly being exposed to geographic information of all kinds via services like Google Maps, Google Earth, and OpenStreetMap, the need for accessible, easy-to-browse geospatial information is as important as ever. The geographic information embedded in map catalogue records is one critical way in which existing records can be integrated into these interactive, browsable systems.

Works Cited

Association of Canadian Map Libraries and Archives. "Core Level Cataloguing for Non-serial Cartographic Materials." January 2008. http://acmla.org/CoreCat_08.pdf (accessed 3 August 2011).

Caldwell, Douglas R. Unlocking the Mysteries of the Bounding Box. August 29, 2005. http://purl.oclc.org.proxy2.lib.uwo.ca:2048/coordinates/a2.htm (accessed May 22, 2011).

Davis, Harry O., and James S. Chervinko. "Map Cataloging and Classification: The Basic Who, What, and Where." In Maps and Related Cartographic Materials: Cataloging, Classification and Bibliographic Control, by Paige G. Andrew and Mary Lynette Larsgaard, 9-37. Binghamton, NY: The Haworth Information Press, 1999.

Donkin, Kate, and Michael F. Goodchild. "A Computerized Approach to Increased Map Library Utility." The Cartographer 4, no. 1 (1967): 39-45.

Ercegovac, Zorana. "Minimal Level Cataloging: What Does It Mean for Maps in the Contexts of Card Catalogs, Online Catalogs, and Digital Libraries?" Journal of the American Society for Information Science 49, no. 8 (1998): 706-719.

Federal Geographic Data Committee. Appendix A: Glossary. February 9, 2006. http://www.fgdc.gov/metadata/csdgm/glossary.html (accessed 24 July 2011).

Gonzalez, Jorge A. Problems That Arise When Providing Geographic Coordinate Information for Cataloged Maps. October 16, 2007. http://purl.oclc.org.proxy2.lib.uwo.ca:2048/coordinates/b8.htm (accessed May 22, 2011).

Goodchild, Michael F. "The Alexandria Digital Library Project." D-Lib Magazine 10, no. 5 (May 2004).

Hill, Linda L., Larry Carver, Mary Larsgaard, Ron Donlin, Terrence R. Smith, James Frew, and Mary-Anna Rae. "Alexandria Digital Library: User Evaluation Studies and System Design." Journal of the American Society for Information Science. 51, no. 3 (2000): 246-259.

Jones, Michael T. "Google's Geospatial Organizaing Principle." IEEE Computer Graphics and Applications. July/August 2007: 8-13.

Larsgaard, Mary Lynette. Map Librarianship: An Introduction. 3rd Edition. Englewood, CO: Libraries Unlimited, 1998.

Larson, Ray R. "Geographic Information Retrieval and Spatial Browsing." In GIS and Libraries: Patrons, Maps and Spatial Information, by Linda Smith and Myke Gluck, 81-124. Urbana-Champaign, IL: University of Illinois, 1996.

Library of Congress. "MARC Proposal No. 2006-DP01." MARC Standards. June 20, 2006. http://www.loc.gov/marc/marbi/2006/2006-06.html (accessed July 14, 2011).

Moore, Susan M., and Lucinda M. Hall. "Map Cataloging: Learning the Basics." San Francisco, CA, 16 June, 2001.

Morris, Steven P. "Geospatial Web Services and Geoarchiving: New Opportunities and Challenges in Geographic Information Services." Library Trends 55, no. 2 (Fall 2006): 285-303.

Petras, Vivien. "Statistical Analysis of Geographic and Language Clues in the MARC Record." Technical Report for Going Places in the Catalog: Improved Geographical Access Project (31 Dec 2004): 1-6.

Riekert, Wolf-Fritz. "Automated Retrieval of Information in the Internet by Using Thesauri and Gazetteers as Knowledge Sources." Journal of Universal Computer Science 8, no. 6 (2002): 581-590.

Rockwell, Ken. "Problem Areas in the Descriptive Cataloging of Sheet Maps." In Maps and Related Cartographic Materials: Cataloging, Classification and Bibliographic Control, by Paige G. Andrew and Mary Lynette Larsgaard, 39-63. Binghamton, NY: The Haworth Information Press, 1999.

Stibbe, H. L. P. "MARC - Maps: The History of its Devleopment and a Current Assessment." Bulletin van de Vakgroep Kartografie no. 3 (Organization for Information Policy, Geografisch Instituut van de Rijksuniversiteit), 1976.

Vestavik, Øyvind. "Geographic Information Retrieval: An Overview." 2004. http://www.idi.ntnu.no/~oyvindve/article.pdf (accessed May 20, 2011).

Webster, G. "Cartographic Materials and MARC." The Cartographic Journal 19, no. 1 (June 1982): 60-67.

Yu, Lixin. "Knowledge Discovery in Spatial Cartographic Information Retrieval." Library Trends 48, no. 1 (Summer 1999): 249-263.

Maps and Inclusion Criteria

Canada Dept. of Mines and Technical Surveys. "Northwest Territories and Yukon Territory". [map] 1:4,000,000. Ottawa: Surveys and Mapping Branch, 1959.

Nunavut was created from part of the former Northwest Territories in 1999.

—. "Yukon Territory" [map]. Edition 2. 1:2,000,000. Ottawa: Surveys and Mapping Branch, 1963.

This map was included since it was found to be present in both collections.

Canada Dept. of Northern Affairs and Natural Resources. "Otto Fiord Glacier 1964. Northern Ellesmere Island, N.W.T., Canada" [map]. 1:50,000. Saint John: University of New Brunswick, 1965.

The location of this feature no longer falls within the boundaries of the Northwest Territories; It has been part of Nunavut since 1999.

Gebrüder Borntraeger Verlagsbuchhandlung. "Afrika-Kartenwerk" [map]. 1:1,000,000. Berlin: Gebrüder Borntraeger, 1976.

This is a German-language map featuring African countries whose names have changed in the years following 1976.

J. H. Colton & Co. "New Brunswick, Nova Scotia, Newfoundland and Prince Edward Id." [map]. 1:3,250,000. New York: J. H. Colton & Co., 1855.

Newfoundland did not become a Canadian province until 1949.

M.S. Boehm & Company Limited. "Map of Busy Berlin" [facsimile]. 1912. 1:5,000. As reproduced by, Ottawa: Association of Canadian Map Libraries and Archives, ACML Facsimile Map Series #130, 1989.

Berlin, Ontario was re-named Kitchener in 1916.

National Geographic Society. "South Asia, with Afghanistan and Burma" [map]. 1:6,522,000. Washington, DC: National Geographic Society, 1984.

Burma was re-named Myanmar by its government in 1989. This decision was highly contested and the name change is not recognized by some countries.

Pathfinder Air Surveys Limited. "Map and Street Guide, Port Arthur, Ontario". [map] 1:16,200. Ottawa: Pathfinder Air Surveys, 1964.

Port Arthur and Fort William amalgamated to form the city of Thunder Bay in 1970.

United States Central Intelligence Agency. “USSR, Travel Restrictions on Foreigners” [map]. 1:12,000,000. Washington, DC: Central Intelligence Agency, 1981.
The Soviet Union was dissolved in 1991.

United States Central Intelligence Agency. “Zaire” [map]. 1:5,000,000. Washington, DC: Central Intelligence Agency, 1973.’
Now known as the Democratic Republic of Congo. Zaire was in use from 1971-1997.

Appendix A. List of terms

AACR – Anglo-American Cataloguing Rules. Developed and published by the American Library Association, the Canadian Library Association, and the UK-based Library Association (now the Chartered Institute of Library and Information Professionals) in 1967. The second edition (AACR2) was published in 1978, and revised in 1988 and 2002.

Bounding box – A typically rectangular "box" enclosing a spatial data set, typically marking its extent. Coordinates are typically provided at each corner to situate the data set in space.

Gazetteer – Geographical index providing, at minimum, a list of place names with their associated coordinate information. Additional information such as feature type is commonly included.

Geolibrary – A library whose search mechanism for all materials (photographs, news stories, music, etc.) is based primarily on geographic location (Goodchild 2004). One early example is the Alexandria Digital Library project at the University of California Santa Barbara.

Geospatial data – According to the US Federal Geographic Data Committee (2006), geospatial data "identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth". This information may include maps, aerial photography, and various forms of remote sensing imagery.

GIS – Geographic Information System. A system that integrates, stores, manipulates, and presents geographically referenced information to inform decision-making.

GPS – Global Positioning System. A global navigation system based on a network of satellites and ground receivers. A user's position on the earth's surface can be calculated by computing the time difference for different satellite signals to reach the user's receiver.

MARC – MAchine Readable Cataloguing. Developed by the Library of Congress in the 1960s, MARC records are a widely adopted standard for the exchange of bibliographic data.

MLC – Minimal Level Cataloguing. Implemented by the Library of Congress in 1978, MLC stipulated the minimum number of data elements that would be required of bibliographic records. It was designed as a cost-effective means to provide access to materials that would otherwise not have received full cataloguing.

Topology – A set of rules governing the spatial relationships between objects. In a GIS, common topological relationships include adjacency, connectivity, and containment.

Appendix B. Relevant geographic MARC fields (Petras 2004, 1-2).

In her 2004 paper "Statistical analysis of geographic and language clues in the MARC record", Vivien Petras summarized the findings of her 5,065,574 record search of the University of California's library catalogue in the following table (10 MARC fields associated with geographic information were selected). From these findings, it appears that almost half of the records have been assigned geographic area codes (043 $a). Geographic subdivision codes and local geographic area codes were significantly rarer.

Field	Field Name	Count	%
008	Control field 008	5065574	100%
$15-17	Place of publication, production or execution		
033	Date/Time and Place of an event		
$b	Geographic classification area code	0	
$c	Geographic classification subarea code	0	
043	Geographic area code		
$a	Geographic area code	2341152	46.22%
$b	Local GAC code	61	
$c	ISO code	0	
044	Country of publishing/producing entity code		
$a	MARC country code	16	
$b	Local subentity code	0	
$c	ISO country code	0	
052	Geographic classification	412	0.008%
260	Publication, distribution, etc. (Imprint)	5048397	99.66%
$a	Place of publication, distribution, etc.		
518	Date/time and place of an event note		
$a	Date/time and place of an event note	4	
522	Geographic coverage note	2	
$a	Geographic coverage note		
650	Subject added entry – topical term		
$c	Location of event	122	
$z	Geographic subdivision	1767693	34.90%
651	Subject added entry – geographic name		
$a	Geographic name	919137	18.14%
$z	Geographic subdivision	42136	0.83%

Appendix C. Example map records for Universities A and B.

A.

◀ Previous   Details: Record 13 of 182   Next ▶

mark request print citation email save mylibrary refworks permanent link

Title   New Brunswick, Nova Scotia, Newfoundland and Prince Edward Id. [map].
J.H. Colton & Co.
Imprint   New York : J. H. Colton & Co., 1855.
1 map : col. ; 29 x 36 cm. Scale [ca. 1:3 250 000] (W 68°--W 52°/N 52°--N 43°30').
Language   English

Other

Holdings | Details | Subjects | MARC View

000			em a0c a
001			2354246
007			jocanzn
008			880729s1855 nyu ea a eng d
034	1		a |b 3250000 |d W0680000 |e W0520000 |f N0520000 |g N0433000
035			(Sirsi) AJZ-2413
039			|f AD
040			CaOTUTF |b eng |c CaOTUTF
052			3415
090	0	9	G 3415 1855 J45 Maps |b RBSC |c 1
110	2		J.H. Colton & Co.
245	1	0	New Brunswick, Nova Scotia, Newfoundland and Prince Edward Id. |h [map].
255			Scale [ca. 1:3 250 000] |c (W 68°--W 52°/N 52°--N 43°30').
260			New York : |b J. H. Colton & Co., |c 1855.
300			1 map : |b col. ; |c 29 x 36 cm.
500			Shows county boundaries, place names; longitude 60 also marked: Longitude East 17 from Washington.
596			32
651		0	Atlantic Provinces |x Maps
948			07/15/1992 |b 06/26/1995

B.

go to result: (end of list)
view: brief full staff (MARC)

000 00681cem a 00217 450
001 2325340
005 20020530070322.0
007 aj ca
008 s1855 nyu 0 eng u
035 __ **|9** CaOGU00738264
052 __ **|a** 3410
099 __ **|a** G 3410. 1855
110 __ **|a** Maritime Provinces
245 10 **|a** New Brunswick, Nova Scotia, Newfoundland and Prince Edward Id. [i.e. Island] **|h** [map].
255 __ **|a** Scale 1:3,379,200.
260 __ **|a** New York : J.H. Colton, 1855.
300 __ **|a** 1 map : col. ; 28 x 36 cm. on sheet 36 x 44 cm.
500 __ **|a** Located in map drawer 11
651 _0 **|a** Maritime Provinces **|x** Maps
651 _0 **|a** Newfoundland and Labrador **|v** Maps.
710 __ **|a** Newfoundland and Labrador.

Appendix D. Data collection instrument

			1		2		3		4		5		6		7		8		9		10	
			Yukon Territory (1963)		Otto Fiord Glacier (1959)		South Asia with Afghanistan and Burma (1984)		USSR, Travel Restrictions on Foreigners (1981)		Map of Busy Berlin (1912) - Facsimile (1989)		Map and Street Guide, Port Arthur, Ontario (1964)		NWT and Yukon Territory (1959)		NB, NS, NL and Prince Edward Id. (1855)		Zaire (1973)		Afrika-Kartenwerk (1976)	
Map location (University)			A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Relevant geographic MARC fields (Petras 2004, 1-2).																						
008		*Control field*																				
	$15-17	**Place of publication, production or execution**	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
033		*Date/time and place of an event*																				
	$b	Geographic classification area code																				
	$c	Geographic classification subarea code																				
043		*Geographic area code*																				
	$a	**Geographic area code**			Y																	
	$b	Local GAC																				
	$c	ISO code																				
044		*Country of publishing/producing entity code*																				
	$a	MARC country code																				
	$b	Local subentity code																				
	$c	ISO country code																				
052		***Geographic classification***	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
260		*Publication, distribution, etc. (imprint)*																				
	$a	**Place of publication, distribution, etc.**	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
518		*Date/time and place of an event note*																				
	$a	Date/time and place of an event note																				
522		*Geographic coverage note*																				
	$a	Geographic coverage note																				
650		*Subject added entry - topical term*																				
	$c	Location of event																				
	$z	**Geographic subdivision**			Y				Y	Y	Y	Y							Y			
651		*Subject added entry - geographic term*																				
	$a	**Geographic name**	Y	Y	Y	Y	Y	Y			Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	$z	Geographic subdivision																				
Geographic coordinate fields																						
034		*Coded Cartographic Mathematical Data*																				
	$a	**Category of scale (NR)**	Y		Y		Y		Y	Y	Y	Y	Y		Y		Y		Y		Y	Y
	$b	**Constant ratio linear horizontal scale (R)**	Y		Y		Y		Y	Y	Y	Y	Y		Y		Y		Y		Y	Y
	$c	Contant ratio linear vertical scale (R)																				
	$d	**Coordinates - westernmost longitude (NR)**	Y				Y		Y	Y		Y			Y		Y					Y
	$e	**Coordinates - easternmost longitude (NR)**	Y				Y		Y	Y		Y			Y		Y					Y
	$f	**Coordinates - northernmost latitude (NR)**	Y				Y		Y	Y		Y			Y		Y					Y
	$g	**Coordinates - southernmost latitude (NR)**	Y				Y		Y	Y		Y			Y		Y					Y
255		*Cartographic mathematical data*																				
	$a	**Statement of scale (NR)**	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	$b	**Statement of projection (NR)**	Y	Y			Y	Y	Y	Y					Y	Y					Y	Y
	$c	**Statement of coordinates (NR)**	Y				Y		Y	Y		Y			Y		Y					Y