

D-Lib Magazine

September/October 2011

Volume 17, Number 9/10

MapRank: Geographical Search for Cartographic Materials in Libraries

Markus Oehrli

Zentralbibliothek Zürich, Maps and Panorama Division, Zürich, Switzerland

markus.oehrli@zb.uzh.ch

Petr Přidal

Klokant Technologies GmbH, Baar, Switzerland and Moravian Library, Brno, Czech Republic

petr.pridal@klokantech.com

Susanne Zollinger

ETH-Bibliothek, Map Library, Zurich, Switzerland

susanne.zollinger@library.ethz.ch

Rosi Siber

EAWAG, Dübendorf, Switzerland

rosi.siber@eawag.ch

doi:10.1045/september2011-oehrli

Abstract

Searches for cartographic materials have previously been carried out using conventional library search systems. However, this kind of search method often proves to be inadequate, as the lack of suitable user interfaces means that queries have to be formulated in words. Furthermore, indexing based on geographical names does not adequately describe the spatial dimension of cartographic material, and geographical names also tend to be ambiguous and inclined to change. A web-based geographical search system (geosearch), which analyzes the geographical coordinates of MARC21 field 034, has been developed for the Swiss research portal for maps. The geosearch enables cartographic material to be found quickly and efficiently on the basis of its location and spatial extent via an intuitively accessible user interface. This is achieved with a cleverly devised ranking algorithm (MapRank[®]) and an innovative indexing mechanism. The number of search results can also be restricted by applying filters relating to the publication period and map scale.

The geosearch is proving to be a powerful search tool which, in an open-ended search scenario, can be used to find roughly double the amount of relevant cartographic material within a short space of time. The system is intuitive to operate and no previous knowledge is required. It also offers further strengths, such as the ability to search through extremely large quantities of data quickly and independently from the various subject heading systems and thesauri. This makes the geosearch an ideal tool for carrying out map searches in metacatalogs.

Keywords: Cartographic material, geographical coordinates, geographical search, geosearch, map library, map portal, MapRank[®], MARC21, mashup, metacatalog, metadata, OPAC, ranking, research portal, spatial indexing, spatial query

Introduction

Conventional library catalogs (OPAC) are only of limited use for map searches. If bibliographic information such as the title or author is not known at the outset, carrying out a text-based search for cartographic material can be a

complicated and time-consuming process. Even the introduction of new filter concepts (faceting) in recent years has done nothing to fundamentally change this unsatisfactory situation.

A typical query contains both topical and geographical terms to establish a spatial reference (e.g. "topography AND St. Petersburg"). The number of search results can then be gradually restricted by applying filters. Completing the steps described, as well as comparing the results of different queries and scanning long lists of hits, requires a disproportionately high amount of time. Judging by our own experience with library patrons over many years, this type of search strategy results in about 20 to 70 percent of the relevant material being found. The percentage depends on the question, previous knowledge and perseverance.

In many respects, the indexing by geographical subject heading (MARC21 field 651) is not particularly suited to making maps easier to find. Geographical names do not always have one distinct meaning. In the German-speaking world, for example, there are at least fifty homonymous places bearing the name of *Neustadt*. Furthermore, due to their cultural imprint, geographical names are not constant ([Buckland et al., 2007](#)). One example of this is the Russian city of *Sankt-Peterburg*, which was known as *Leningrad* from 1924 to 1991. Thirdly, geographical names are language-specific: in search queries, the endonym *Санкт-Петербург* (Sankt-Peterburg) is also used exonymically in variations such as *Saint Petersburg*, *Saint Pétersbourg*, *Sankt Petersburg*, etc. Finally, usually only one to three of the most relevant geographical subject headings are recorded by the library for each document, but this only enables a geographical area to be described in a very generalized way. In the majority of cases, a search for any given place charted in the cartographic material therefore yields no results.

By contrast, geographical coordinates, which specify the location and spatial extent of a map precisely, do not present any of these disadvantages. The geographical coordinates allocated to a particular point on earth are unique and, for the purposes of library systems, constant. Inconsistencies may occur, within an acceptable range of dozens of meters, because the geodetic datum and coordinate reference system on which the document is based are not taken into account when compiling the metadata. The [MARC21](#) standard stipulates fields 034 and 255 for recording geographical coordinates and scale.

001				vtls001404246
003				Sz
005				20110318032700.0
008				041011s2004 sz abg a 0 eng d
020				\$a 3906593355 : \$c CHF 22.80
034	1			\$a a \$b 500000 \$d E0703000 \$e E0760000 \$f N0393700 \$g N0364000
035				\$a Sz001404246
040				\$a Sz \$c Sz
245	0	0		\$a The Pamirs \$h [Kartenmaterial] : \$b a tourist map of Gorno Badakhshan - Tajikistan and background information on the region / \$c Markus Hauser
246	3	0		\$a Pamirs \$h [Kartenmaterial]
255				\$a 1:500 000 \$c (E 70°30'--E 76°00'/N 39°37'--N 36°40')
260				\$a Hinteregg (8132) : \$b distrib.: Gecko Maps, A. Rohweder, \$c cop. 2004
300				\$a 1 Karte : \$b mehrfarbig ; \$c 64 x 94 cm, gef. 17 x 11 cm
500				\$a 3 Nebenkarten auf Rückseite: Übersichtskarte ; Khorog 1:30 000 ; Murghab 1:15 000
651		7		\$a Pamir \$v Karte \$2 SWD
700	1			\$a Hauser, Markus

Fig. 1: Structure of a MARC21 record: the fields 034 and 255 contain the scale and the geographical coordinates (highlighted in red). Source: Swiss National Library.

Choosing to include geographical coordinates in library records is, according to recent literature, regarded as one of the most important decisions for improving the accessibility of cartographic material ([Bidney, 2010](#)). This naturally calls for systems that take geographical coordinates into account in search queries.

Previous geosearch approaches and solutions in libraries

In Switzerland, the idea of making maps easier to find with the help of geographical coordinates has been discussed for decades. As early as 1986, at the instigation of the ETH-Bibliothek, a translated Swiss version of the ISBD(CM) noted that the specification of coordinates would be advisable when using electronic data processing ([Katalogisierungsregeln](#), 1986). The recording of coordinates was made obligatory in 1995, as stated in an internal version of the regulations of the ETH-Bibliothek and Zentralbibliothek Zürich. This rule was retained when the [AACR2](#) was adopted at national level in 1999 even though the AACR2 at the time still referred to the recording of geographical coordinates as an "optional addition" ([AACR2](#), 1998). The ETH-Bibliothek was unable to put detailed plans for a web-based geosearch for single maps and map series into practice in 2004 due to the high licensing and development costs involved ([Bühler](#), 2004).

The successes of Google Maps and Google Earth, as well as the development of new web technologies, the availability of open application programming interfaces (API) and the dynamic progress of the open source movement ([Open Source Geospatial Foundation](#)), have transformed the development of web-based user interfaces with map components into a simple and cost-effective process. Moreover, open source components enable the flexible adaption of applications to exploit new technological possibilities ([Schüler and Crom](#), 2008).

The number of geosearch applications for map libraries available on the Internet is growing rapidly. These mostly consist of mashups that use the Google Maps API (e.g. [Maps of Australia](#); [Maps of North Carolina](#); [MapFast](#); [Geographische Kartenblattsuche](#); [MapHappy](#); cf. [Johnston and Jensen](#), 2009). These applications are interesting prototypes, but they are designed purely for handling searches in fairly small collections of data. Furthermore, none of the applications mentioned above locate all of the cartographic material that is contained within a given query region or that touches on or overlaps with it. To reduce the complexity of the projects, searches are often based on points (the central or corner points of a map), although a search by area is preferred.

The installation and configuration of a geosearch for map libraries that is based on open source software was, in addition to the development of other applications, the subject of the [OldMapsOnline](#) project, which was conducted during the period from 2008 to 2011 by the Moravian Library in Brno, Czech Republic. This project involved testing applications including [TimeMap](#), [GeoNetwork](#), [Lucene](#) and [LuceneGeoTemporal Extensions](#) from the DigMap project, [PostGIS](#), [Alexandria Digital Library](#) and the proprietary online service [Google Geo Search](#). However, the test results revealed that none of these applications was able to provide an easy-to-operate geosearch interface. In addition, none of the applications offered an adequate ranking system for large collections of data.

Requirements for a modern geosearch application

In 2008, the opportunity arose in Switzerland to develop a national research portal for maps ([Kartenportal.CH](#)) as part of the Swiss electronic library collaborative project [e-lib.ch](#) ([Klöti and Schmid](#), 2011). The following institutions were involved in the project: the ETH-Bibliothek, the WSL and EAWAG-EMPA research institutions, Zentralbibliothek Zürich, the Swiss National Library, the University of Bern library, and the Institute of Cartography at ETH Zürich. The most important and most challenging sub-project was the implementation of a geosearch.

In parallel to the national map portal – and also within the scope of the e-lib.ch project – the [swissbib](#) metacatalog was established. Swissbib covers approximately 15.3 million documents from 730 Swiss libraries and archives under a single interface, including approximately 130,000 records of cartographic material (as of July 2011). This presented major advantages for the planned geosearch: the laborious process of harvesting and deduplicating the MARC21 metadata would no longer be required, as the map portal would access the processed metadata via an [OAI-PMH](#) interface directly from the metacatalog. In addition, it would be possible to display the detailed metadata and holding records of search results from the geosearch via a link back to the metacatalog.

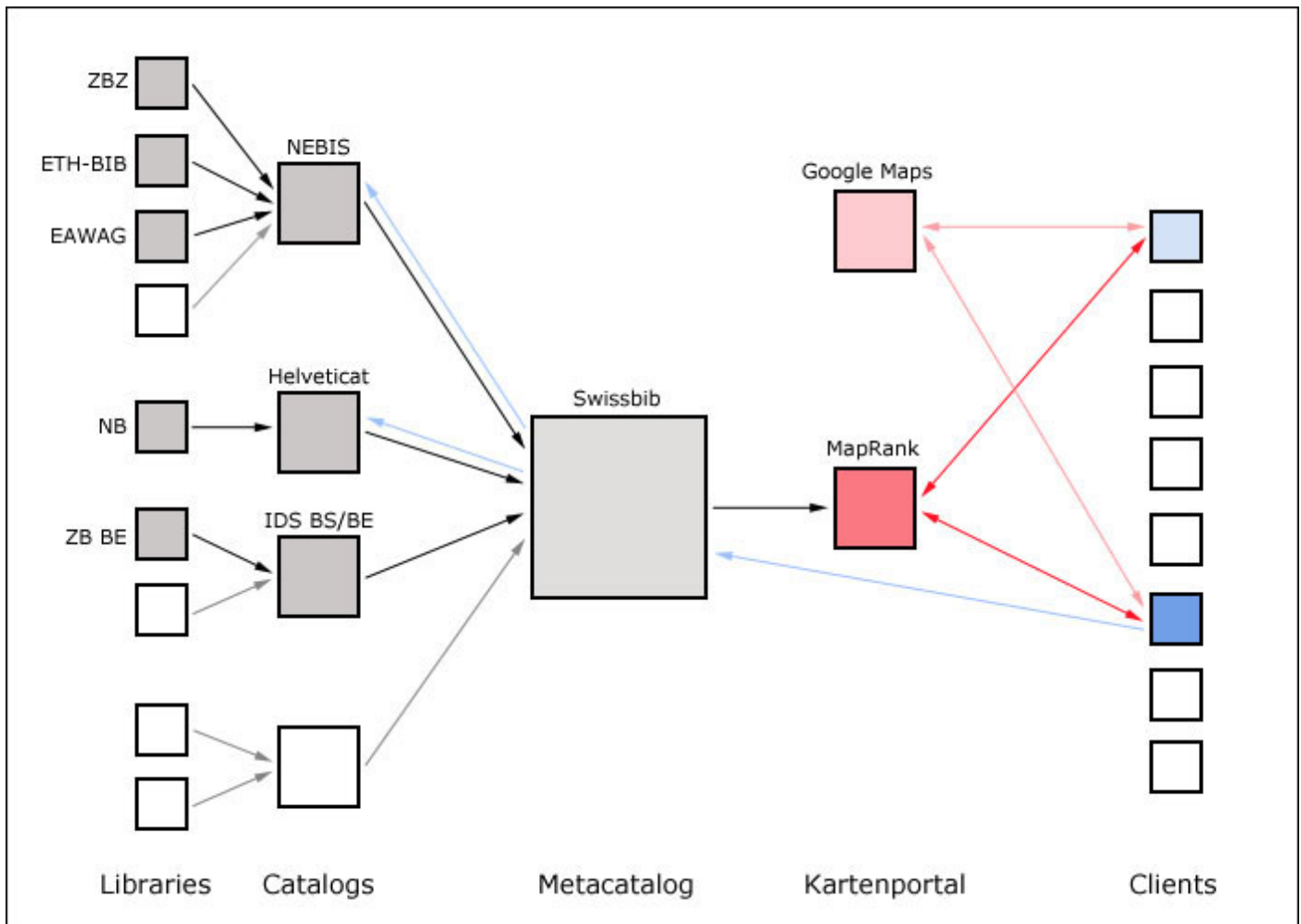


Fig. 2: Data flow (from left to right): metadata are captured by the libraries and managed in the local catalogs. The metadata are collected and deduplicated in the national metacatalog. From here the records are transferred into the MapRank geosearch. The geosearch interface runs in the user's web browser (red arrows). Detailed metadata from the metacatalog are displayed; the user is referred back to the local catalogs to order the document (blue arrows). Diagram: authors.

All in all, the complexity of the geosearch sub-project was reduced to such an extent that it could be implemented using the resources available.

The following key requirements were defined for the planned geosearch:

- The user interface must be user-friendly.
- The transfer of coordinates to the search system must be carried out intuitively by defining a rectangle in a reference map from an online map service (hereafter "search map").
- The search map must cover the entire world with a sufficient level of detail.
- The search results must cover all maps contained within the search rectangle as well as those that overlap with this area (area search).
- It must be possible to restrict the number of search results by applying optional filters relating to the publication period and scale.
- It must also be possible to search through very large collections of data, containing up to 500,000 records, without any loss of speed.
- The system must offer quick response times and therefore a high interaction speed.
- An integrated text-based search is considered to be of secondary importance.
- If possible, open source components must be used.

The fact that a complete list of search results for each place on earth covers a range extending from large-scale maps up to all world maps meant that it was essential for the ranking of the results to prioritize relevant documents.

As none of the existing applications was capable of fulfilling all these requirements, a public invitation to tender was issued and the company [Klokian Technologies](#) was subsequently commissioned to develop an innovative solution.

Current technologies for processing spatial queries

The problem of searching by geographical data (geodata) is a well-known issue from programming geographic information systems (GIS). Spatial databases such as PostGIS, SpatiaLite or Oracle with extensions offer functionality for handling spatial queries. The records are usually indexed with variations of an R-tree or a Quadtree. The algorithms are able to efficiently answer queries such as: "Which points are located within the query region?" or "Which polygons overlap with a given query region?".

In geodata catalogs, searches are carried out for data that are represented by bounding boxes. One example of this is [GeoNetwork](#). However, these systems are also often designed for searching in relatively small collections of data and they frequently generate very long lists of search results that are not sorted according to relevance. Users must apply filters to limit the number of displayed documents. The remaining results have to be scanned by the user, record by record, to find relevant information. It is possible to enhance these systems with a ranking function that works on the basis of a spatial similarity function, such as the Hausdorff distance or the ratio between the size of the intersection area and the query region. Such a similarity function is used to sort the results from the spatial query before they are displayed to the user. However, this kind of method is not quick enough for searching large databases, as the number of records is too high for a real-time ranking calculation.

Another approach is to use a form of indexing that is commonly found in full text databases: an inverted index. This allows existing full text search engines, such as [Lucene](#), to be used to index geographical records. This method – excluding the ranking element – is described under the name of C-squares ([Rees, 2003](#)). If this indexing system is applied, the same grid (sometimes known as the Spherical Mercator) can be used as in Google Maps, OpenStreetMap and other well known map services ([Přidal, 2008](#)). Lucene has its own extension for spatiotemporal searching (LGTE), but it offers limited potential because the records are represented as points and the ranking relies only on simple distances.

As an alternative to encoding geographical coordinates and implementing a simple geosearch, some search systems use [GeoHash](#) or a variation of this. GeoHash encodes pairs of longitudes and latitudes as a string of characters that feature interesting characteristics: places located close to each other have similar prefixes. Characters can be cut from the end of the string, thus making the string shorter, but also less precise. The same principle can also be applied to encoding regions using bounding boxes. GeoHash uses interleaved bits for encoding a traversing path in the Euclidean space. However, the algorithm fails in edge cases, which is why workarounds have been created. This type of text string, together with the alphabetical sorting system that is implemented in every database, can be used for a simple geographical proximity search.

To summarize, the review of the current technologies exposed some shortcomings that called for enhancements to be developed.

Technical implementation using MapRank

Klokian Technologies developed a new ranking algorithm known as MapRank[®]. MapRank is based on an extended spatial similarity function, i.e. a function that comprises two regions: the region covered by the cartographic material and the region covered by the user's search query. These two regions are typically linked by one of the following spatial relationships: equals, overlaps, contains, is contained by, disjointed ([Larson and Frontiera, 2004](#)).

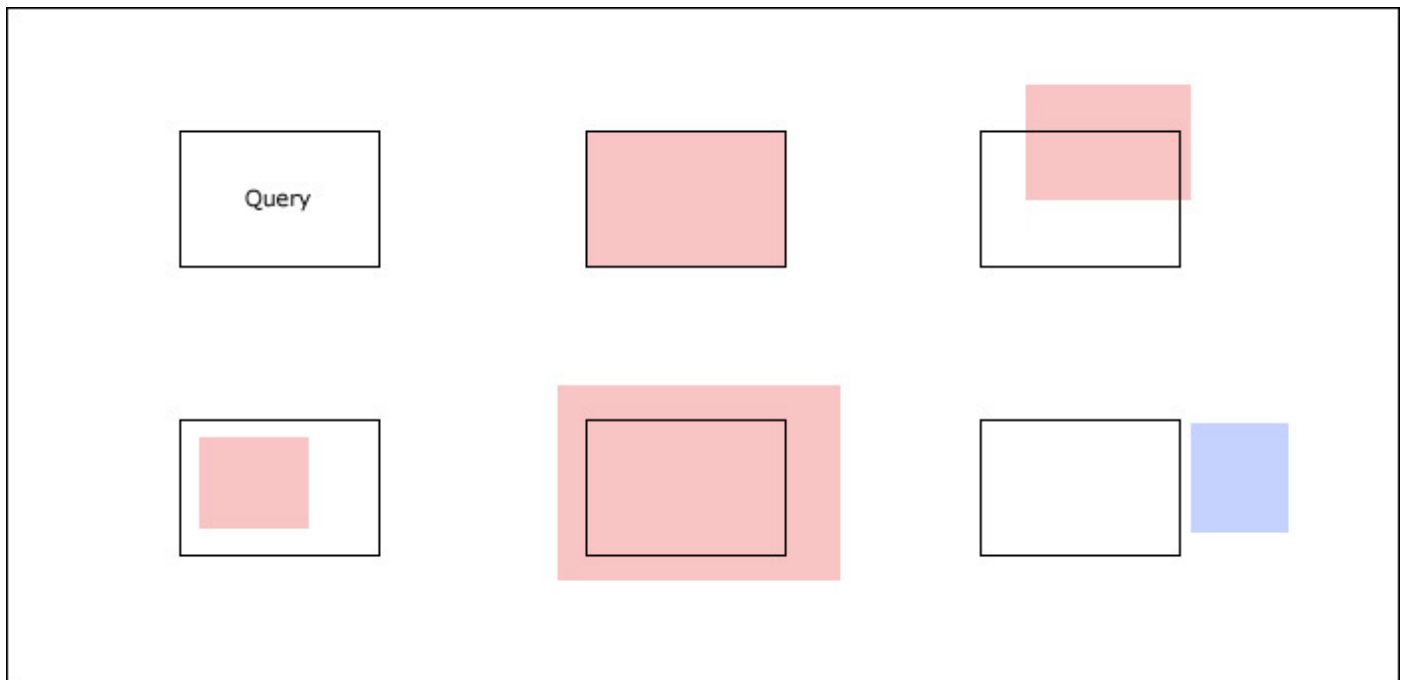


Fig. 3: Principle of the spatial query: some records (red) are taken into account as search results for the query (black); one record (blue) is not considered. Diagram: authors.

The result of the similarity function is combined with additional user input (filters relating to publication period and scale, and full text search) to generate the final ranking, with the most relevant document placed at the top of the list of results. This is therefore in line with the probability ranking principle, whereby the documents are sorted in descending order according to their assumed probability of relevance for the user ([Robertson, 1977](#)).

MapRank is based on the idea of creating an algorithmic description of what is intuitively regarded as a spatial similarity in terms of size, shape and position. The system performs a coverage similarity function, which answers the following question: "which document covers a similar area to the user query defined in the search map?".

Another important element of the geosearch with MapRank is the new type of indexing involved. As has already been mentioned, conventional indexing with an R-tree does not perform well. MapRank, however, uses an index that already contains part of the precalculated ranking of the documents. The precalculation is carried out when the records are imported into the system. This therefore makes the system faster and more scalable, as the speed with which the list of results is generated is not significantly affected by the number of indexed records.

The search interface was written in JavaScript, using Google's Closure Library. The search map and the directory of places are integrated via the [Google Maps API](#). If required, this can be substituted with OpenStreetMap components. The server-side element of the geosearch consists of a FastCGI module, written in C/C++, that runs on a Linux system using a lighttpd web server. The web server supplies the data in JSON format, with the search parameters in line with the [OpenSearch](#) standard wherever possible. With the exception of the MapRank algorithm, open source components are used throughout.

Functionality of the MapRank geosearch

The [first MapRank geosearch](#) version was activated in September 2010. The geosearch interface runs in a standard web browser and does not require any additional software on the client side. It is divided into three parts:

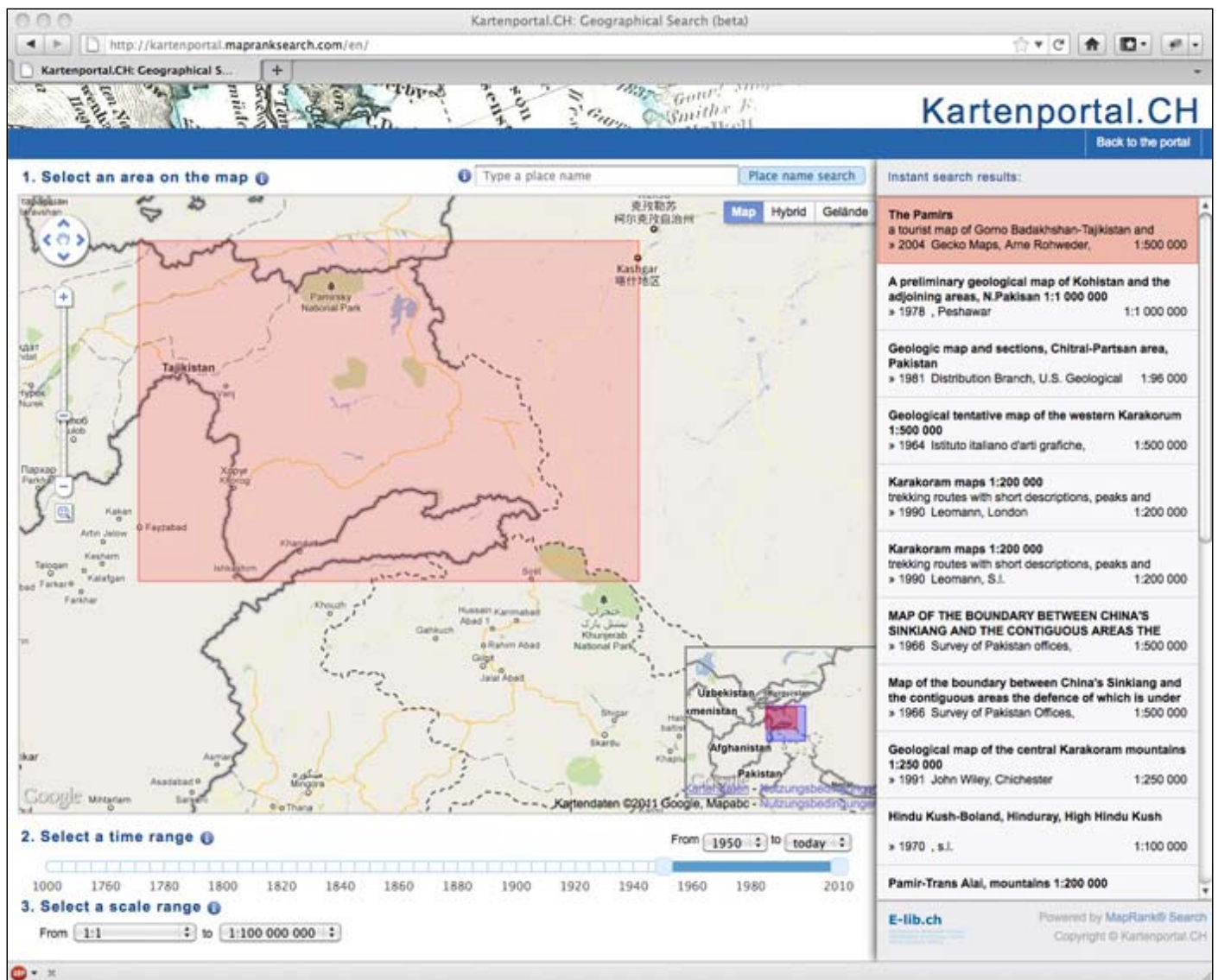


Fig. 4: MapRank geosearch user interface: the most relevant search result is highlighted in red in the results list and on the search map. Source: Klokian Technologies with map data from Google Maps.

1. The most eye-catching feature is the search map from Google Maps. The standard Google Maps zoom and drag function is familiar to users and adds to the convenience of using the geosearch. Alternatively, a geographical name can also be entered and the search map is then centered on this name. The location and spatial extent of a search result is shown on the search map in the form of a red rectangle.
2. Optional filters are included below the search map. The feature that stands out here is the timeline for specifying the publication period, which is displayed on a scale from the year 1000 to 2010. The time range that is most frequently used in searches, 1950 to 2010, is set as default. If necessary, users can set a different time range using two scroll bars or a selection list. The filter for the map scale can be activated via a selection list. The default setting ranges from 1:1 to 1:100,000,000, which covers all of the cartographic material which may be detected in a search.
3. The column on the right-hand side shows the search results in text form. This comprises brief bibliographic information, including the title, subtitle, year of publication, publisher, place of publication and scale. The most relevant results are shown at the top of the list and highlighted in red, corresponding with the red rectangle on the search map. By simply hovering over the search results list with the mouse, it is possible to highlight another result in the list and, in parallel, on the search map. This helps to avoid numerous visualization problems (Janée, 2006).

All that is required for a search query is to center the search map on the desired location and select a zoom level that displays the targeted region with the greatest possible accuracy. Each modification of the search map or the

filters immediately triggers a recalculation of the results list. Clicking on the title in the list that is highlighted in red opens a new window displaying the complete bibliographic metadata contained in the metacatalog, which also indicates and provides links to the holding records.

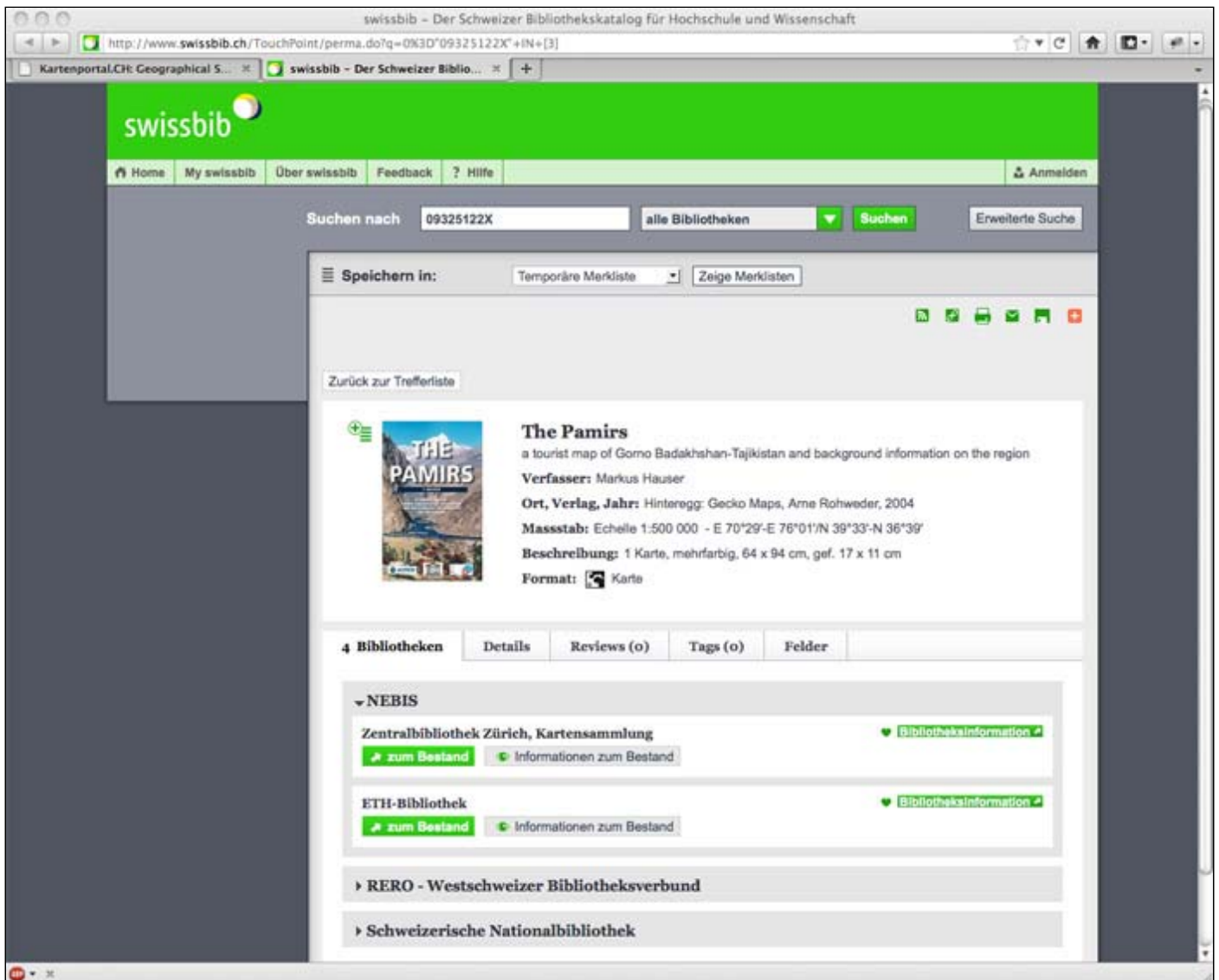


Fig. 5: Metacatalog user interface: when the search result marked in red or the red rectangle is clicked in the MapRank geosearch, the metacatalog user interface is opened in a new window for the purpose of displaying the detailed metadata. The maps that have been found can now be ordered, if the user wishes, from one of the local catalogs. Source: swissbib.

When searching for old maps, the time range for the search must be adjusted on the timeline so that only maps with relevant years of publication (MARC21 field 008) are shown. Alternatively, and with another purpose in mind, one can also consider the analysis of chronological subject headings (MARC21 field 648). However, this feature has not been implemented in the current geosearch installation.

The scale filter is required when searching for a map series should those sheets not have been recorded individually in the metacatalog. The MapRank algorithm specifically prioritizes search results that have approximately the same spatial extent as the specified region on the search map in accordance with the concept of coverage similarity. Preference is therefore given to single maps, while large map series are moved much further down the results list. The scale filter can be adjusted to ensure that map series are displayed.

Initial experiences

The initial experiences of using the geosearch have been consistently positive. Users are generally surprised to be shown search results right from the start, even though they have not knowingly prompted any action. The

navigation system in the widely used Google Maps search map is familiar to many users. Every action carried out by the user on the search map immediately gives rise to visible changes in the ranking of the hits in the search results display and thus quickly achieves a successful outcome. The timeline, which users have found to be easier to operate than the new filter concepts used in conventional OPACs (faceting), met with an enthusiastic response.

The geosearch has proven to be very suitable for search scenarios where users cannot provide any bibliographic data for the cartographical material they are searching for or where the search is very open-ended. A considerable amount of time can be saved in these cases. However, if the precise bibliographic data are known, a text-based search in a conventional OPAC is more efficient and therefore remains the preferred option. Judging by our experience, in the case of open-ended searches, the new search strategy enables users to find 70 to 95 percent of the relevant cartographic material within a short space of time and without any previous knowledge. This also includes documents that have incomplete or, from the user's point of view, unexpected indexing, making them unlikely to be found in a text-based OPAC. This percentage estimate is based on the assumption that the metadata of all map materials of the library includes coordinates and map scale.

The geosearch fully demonstrates its strengths when it comes to searching data in a metacatalog. While word-based search queries in metacatalogs often produce incomplete lists of results when different indexing systems and thesauri clash with each other, the results generated by the geosearch are complete and comprehensible.

However, many records occurred multiple times in the geosearch because they had not been identified as duplicates in the metacatalog. To ensure that records are deduplicated, it is absolutely essential that the libraries involved, as authors of the records, interpret and apply the cataloging rules (in this case, AACR2) and the metadata format (in this case, MARC21) thoroughly and consistently.

From a cartographic perspective, the results display of world maps in the search map is not yet perfected. The search for cartographic material in polar regions has also proven to be unsatisfactory, because, for reasons of projection, Google Maps cannot offer a normal view of the globe. The name search used by Google Maps, which still seems to offer scope for development with regard to homonyms and exonyms, is beyond the sphere of influence of the geosearch project.

Outlook

Plans are being made to optimize the list of search results. This would involve displaying different editions of the same cartographic material, which currently appear in various places in the results list, chronologically and in a condensed format. Furthermore, the intention is for records to be loaded from the metacatalog periodically in future via OAI-PMH interfaces.

The first geosearch version, released in September 2010, contained approximately 68,500 records with correctly documented geographical coordinates. Approximately 56,400 further records did not include any geographical coordinates, or only contained deficient coordinates. Over the next two years, the libraries involved in the project will be focusing on rectifying these deficiencies and retrospectively recording the geographical coordinates. Among other material, this affects maps published before 1800, as until now the documenting of geographical coordinates for old maps was still optional. However, the geosearch is particularly useful when it comes to searching for old maps, which, due to drastic changes in political borders and a frequent lack of bibliographic information, are often very difficult to find.

The ETH-Bibliothek Map Library is planning to import around 250,000 records of sheets from map series into the regular library catalog, which have been recorded in the proprietary local application Toporama up to now. In so doing these records will become available, and therefore easy to find, in the geosearch like the records of single maps.

In future, it should be possible to only display cartographic material from selected libraries by applying an additional filter. Implementing this filter function enables the geosearch to be integrated into the website of each of the libraries involved in the project, to carry out searches exclusively in its own holdings. Moreover, it would be possible to implement filters to enable searches to be performed specifically for atlases or remote sensing images, for example, or with a text search for certain authors or publishers. The [second MapRank geosearch](#) version that has been set up since June 2011 for the David Rumsey Map Collection shows the possibility of integrating a text search and thumbnails, as does the [third MapRank geosearch](#) version for the Moll Collection of the Moravian Library.

It would be possible to make all library holdings having spatial extent accessible in the geosearch as long as the metadata had been supplemented with geographical coordinates, which is not currently envisaged by any of the known sets of cataloging rules. This would offer tremendous potential, including for publications such as artistic works (e.g. *Vue de la chute du Rhin*), travel guides (e.g. *Baedeker's Switzerland*), nonfiction books (e.g. *The Battle of Stalingrad*), works of fiction (e.g. *Hamlet*), movies (e.g. *An American in Paris*), and sound recordings (e.g. *Il barbiere di Siviglia*), etc. In the long term, the geosearch would take its place alongside the text-based OPAC as an equally useful solution for conducting searches in library holdings of all kinds.

Conclusions

The MapRank geosearch presented here is proving to be a powerful search tool. Used to supplement the text-based OPAC, it makes it considerably easier to find cartographic material in libraries as well as saving a significant amount of time. The MapRank geosearch fully demonstrates its strengths when it is used to search metacatalogs, such as its ability to perform an efficient search in extremely large quantities of data, its robustness in handling data from different origins and its independence from any indexing systems. The recording of geographical coordinates has gradually become an established practice in Switzerland since 1986 and this is considered a key factor in the success of the geosearch in the national map portal.

Acknowledgements

The authors would like to thank Jost Schmid, Walter Raaflaub and Vaclav Klusák. This paper has been translated with funds provided by various institutions.

Literature

- [1] AACR2: Anglo-American cataloguing rules. 2nd ed., 1998 revision. Ottawa: Canadian Library Association, 1998.
- [2] Bidney, Marcy M.: Can Geographic Coordinates in the Catalog Record Be Useful? In: *Journal of Map & Geography Libraries* 6, 2 (2010) pp. 140–150. Online: <http://www.tandfonline.com/doi/abs/10.1080/15420353.2010.492304>.
- [3] Buckland, Michael ... [et al.]: Geographic Search: Catalogs, Gazetteers, and Maps. In: *College & Research Libraries* 68, 5 (2007) pp. 376–387. Online: <http://crl.acrl.org/content/68/5/376>.
- [4] Bühler, Jürg: Kartenkataloge der Zukunft: die räumliche Suche in einem graphischen Katalog. In: Bühler, Jürg; Zögner, Lothar (eds.): *Die digitale Kartenbibliothek: eine Momentaufnahme*. München: Saur, 2004. (Kartensammlung und Kartendokumentation/Bibliographia Cartographica. Beiheft; 1). pp. 215–221.
- [5] Janée, Greg: Spatial Footprint Visualization. 2006. Online: <http://www.alexandria.ucsb.edu/archive/2006/footprint-visualization/>.
- [6] Johnston, Lisa R.; Jensen, Kristi L.: MapHappy: A User-Centered Interface to Library Map Collections Via a Google Maps “Mashup”. In: *Journal of Map & Geography Libraries* 5, 2 (2009) pp. 114–130. Online: <http://purl.umn.edu/92083>.
- [7] Katalogisierungsregeln: internationale standardisierte bibliographische Beschreibung für Kartenmaterialien: deutsche Fassung der englischen Originalausgabe, IFLA 1977 [...] 2., überarbeitete Aufl. Bern: Vereinigung Schweizerischer Bibliothekare, 1986. (Katalogisierungsregeln; Fasz. BE).
- [8] Klöti, Thomas; Schmid, Jost: Suche nach gedruckten und digitalen Karten mit Kartenportal.CH. In: *Arbido* 3 (2011) [in print].
- [9] Larson, Ray R.; Frontiera, Patricia: Spatial Ranking Methods for Geographic Information Retrieval (GIR) in Digital Libraries. In: Heery, Rachel ... [et al.] (eds.): *Research and advanced technology for digital libraries: 8th European conference: proceedings ECDL 2004, Bath, UK, September 12–17, 2004*. Berlin: Springer, 2004. (Lecture notes in computer science; 3232). pp. 45–56. Online: http://dx.doi.org/10.1007/978-3-540-30230-8_5.
- [10] Přidal, Petr: Tiles à la Google Maps: Coordinates, Tile Bounds and Projection. 2008. Online: <http://www.maptiler.org/google-maps-coordinates-tile-bounds-projection/>.

- [11] Rees, Tony: "C-Squares": a New Spatial Indexing System and its Applicability to the Description of Oceanographic Datasets. In: *Oceanography* 16, 1 (2003) pp. 11–19. Online: http://www.tos.org/oceanography/issues/issue_archive/issue_pdfs/16_1/16.1_rees.pdf.
- [12] Robertson, S.E.: The probability ranking principle in IR. In: *Journal of Documentation* 33, 4 (1977) pp. 294–304. Online: <http://dx.doi.org/10.1108/eb026647>.
- [13] Schüler, Mechthild; Crom, Wolfgang: GOKaRT: Graphical Online Search Tool for Maps. In: *LIBER Quarterly* 18, 2 (2008) pp. 299–303. Online: <http://liber.library.uu.nl/publish/articles/000252/index.html>.
-

Additional Material

A video demonstrating the MapRank geosearch is online at <http://www.klokantech.com/mapranksearch/>.

Glossary

Bounding box: a model for specifying the geographical coordinates in the metadata for cartographic material. According to AACR2, this should comprise the most westerly and most easterly longitudes and the most northerly and most southerly latitudes of the displayed region.

Map series: cartographic document containing multiple sheets, covering an area that is too large to be displayed on a single map sheet of the same scale.

MapRank®: an algorithm for ranking cartographic material in the geosearch. MapRank concentrates on analyzing the geographical coordinates in the metadata. Produced by Klokant Technologies in Baar, Switzerland.

R-tree: a data structure that enables searches to be carried out for large, multidimensional items, such as points and polygons in a two-dimensional space.

About the Authors

Markus Oehrli is a cartographer and map librarian. He started his professional cartography training at the Swiss Federal Office of Topography. Since 2003, he has been the rare maps and map series cataloger at the Maps and Panorama Division of the Zentralbibliothek Zürich, Switzerland. Markus served for almost ten years on the board of editors of *Cartographica Helvetica*, the German-speaking journal on the history of cartography.

Petr Přidal is a software engineer. Since 2008 he has been working on the OldMapsOnline.org and TEMAP research projects at the Moravian Library in Brno, Czech Republic. This resulted in the improvement of several open source software tools (such as IIPImage JPEG2000, MapTiler, OpenLayers Zoomify, and GeoTools). Within these projects, he has prepared an open-source-based workflow for processing and the online publishing of scanned maps in public libraries. Petr is the founder and CEO of Klokant Technologies in Baar, Switzerland. The company's innovative software tools for libraries and cultural heritage institutions include MapRank geosearch, Georeferencer, and Geoparser.

Susanne Zollinger is a geographer and GIS specialist. She works in the map collection of the ETH-Bibliothek in Zürich, Switzerland. Susanne is responsible for the introduction of new methods and technologies that improve access to printed and digital maps.

Rosi Siber is a GIS specialist and biologist. She started her work at EAWAG (the aquatic research institute within the ETH Domain, Zürich, Switzerland) in 2002. She is responsible for the GIS support and the geodata management. Rosi was the project manager for the geosearch described here.

Copyright © 2011 Markus Oehrli, Petr Přidal, Susanne Zollinger and Rosi Siber

