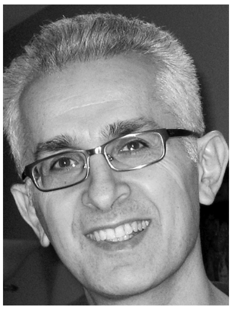
Extending Map-Based Visualizations to Support Visual Tasks: The Role of Ontological Properties†

Olha Buchel* and Kamran Sedig**

*Faculty of Information and Media Studies, The University of Western Ontario, 1151 Richmond Street, London N6A 5B7 ON, Canada, <oburchel@uwo.ca>

**Faculty of Information and Media Studies & Department of Computer Science, The University of Western Ontario, 1151 Richmond Street, London N6A 5B7 ON, Canada <sedig@uwo.ca>

Olha Buchel is a Ph.D. candidate in Library and Information Science at the Faculty of Information and Media Studies, University of Western Ontario, Canada. She received her M.Sc. in LIS from the University of Illinois at Urbana-Champaign. Previously, she was a member of the research team at the University of California at Santa Barbara, where she was involved in the design of metadata and classroom applications for digital libraries. Her research interests include map-based visualizations, computer-mediated human-information interaction, visual tasks, and sensemaking activities.



Kamran Sedig is an Associate Professor at the University of Western Ontario in Canada. He has a joint position in the Faculty of Information and Media Studies and the Department of Computer Science. He received his B.Sc., M.Sc., and Ph.D. from Concordia, McGill, and the University of British Columbia, respectively. His research interests include computer-mediated human-information interaction, design of interactive information artifacts, and cognitive and knowledge technologies.

Buchel, Olha and Sedig, Kamran. **Extending Map-Based Visualizations to Support Visual Tasks: The Role of Ontological Properties.** *Knowledge Organization, 38(3)*, 204-229. 117 references.

ABSTRACT: Map-based visualizations of document collections have become popular in recent times. However, most of these visualizations emphasize only geospatial properties of objects, leaving out other ontological properties. In this paper we propose to extend these visualizations to include non-geospatial properties of documents to support users with elementary and synoptic visual tasks. More specifically, additional suitable representations that can enhance the utility of map-based visualizations are discussed. To demonstrate the utility of the proposed solution, we have developed a prototype map-based visualization system using Google Maps (GM), which demonstrates how additional representations can be beneficial.

Received 14 October 2009; Revised 18 December 2009; Accepted 2 January 2010

† The authors would like to thank the reviewers for their valuable comments to improve the quality of the paper. We would also like to thank Paul Parsons for his assistance to improve the quality of the paper.

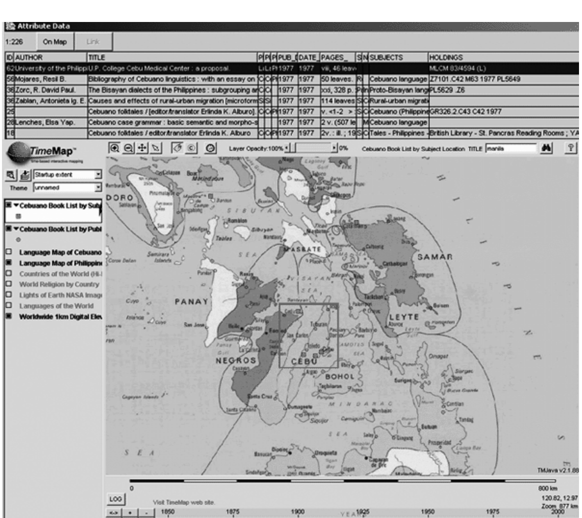
1.0 Introduction

In recent times, visualizing geospatial properties of books and other documents has become increasingly popular. Examples of such visualizations are book rings in Google Earth (Jones 2008, 29min. 30 sec.), visualizations of geographical references illustrating Google Books Search (http://books.google.com/), Atlas of Fiction (http://www.atlasoffiction.com/), Bible Map (http://www.biblemap.org/), and many others. These visualizations appeal to users due to their rich interaction possibilities (i.e., zooming, panning, flybys) and innovative and interesting ways of presenting and accessing information about documents.

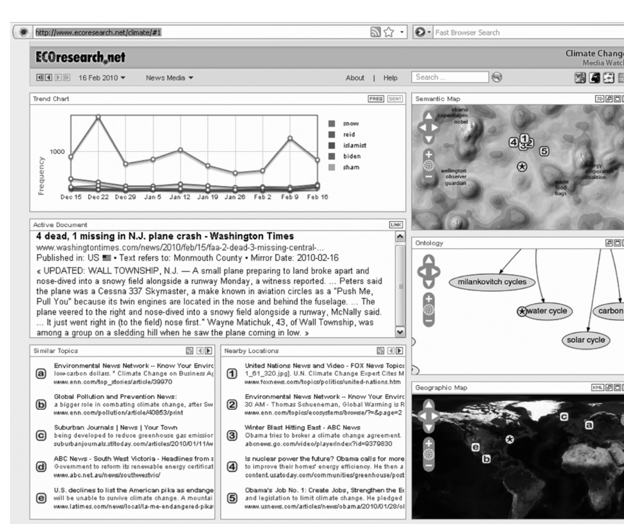
Significant theoretical work has been done for the visualization of documents on digital maps in digital library projects. A project of Electronic Cultural Atlas Initiative (ECAI) used a digital map to facilitate searching a collection of 700 MARC records about, or published in, the Cebuano region of the Philippines (Buckland et al. 2007). Another project, *Going Places in the Catalog: Improved Geographic Access* (Buckland et al. 2004), has experimented with the translation of spatial queries, drawn on a map in various graphical and textual forms, and time/space visualizations of documents using the TimeMap software (Archaeological Computing Laboratory, University of Sydney 2004).

Geovisualization researchers have also been interested in the visualization of documents. Stryker et al. (2008) describe a map-based visualization for tracking infectious disease threats such as Avian Influenza in the media. In their visualization, text extracted from documents can be queried and represented as visual artifacts within a map, timeline, or an extended tag cloud. These linked representations enable the user to progressively filter a collection of documents by interacting with location, time, and theme. A slightly different set of coordinated representations is implemented in the Media Watch on Climate Change visualization (a domain-specific news aggregation portal available at http://www.ecoresearch.net/climate/) (Hubmann-Haidvogel et al. 2009). Representations in this visualization include a synchronized geospatial map, a set of three semantic views (Semantic Map in a form of a spatialization, Ontology View, where concepts are interrelated, and Tag Cloud), and a linear frequency graph of the five most popular keywords. Views represent the different dimensions of the contextualized information spaces, providing the user with multiple perspectives on the latest news media coverage.

Regardless of who develops such visualizations, they have similar designs, mainly consisting of maps, timelines, small legends, and representations of themes and subjects. Such designs are space-, time-, and topic-centred; that is, their primary focus is the visualization of geographic space, time, and sometimes topics or themes. Below we present two examples of map-based visualizations which demonstrate this approach. One is the TimeMap software (see Figure 1a), which is often used for visualizing library collections, and the other is the Climate Change Media Watch, which is used by geoscientists for interpreting news about climate research. Both visualizations have maps, timelines, and representations of subjects (for example, the TimeMap has a drop down list of themes; and the Climate Change Media Watch has three views: the Semantic Map, the Ontology View, and the geospatial map which alternates with the Tag Cloud).



a)



b)

Figure 1. Examples of space-time-and-topic-centric map-based visualizations: a) TimeMap interface (Zerneke et al. 2006); b) Climate Change Media Watch Interface (which can be accessed at http://www.ecoresearch.net/climate/).

The noticeable elements absent in these visualizations are the graphical representations of ontological properties of documents. In terms of library and information science, ontological properties are resource descriptors that describe properties of documents, not information about metadata records, metadata contributors, or metadata creators. Ontological properties of documents may include information about title, various kinds of subjects, authorship, explicit relationships among documents, coverage, rights, language, form/genre, illustrative matter, bibliography, time of acquisition, time of publication, classifications, citations, and others. Some of these other properties have been visualized in Envision Digital Library Project (1991-1995) (Fox et al. 1993; Fox et al.1995; Heath et al. 1995; Nowell 1997) and in Film-Finder (Ahlberg and Shneiderman 1994; Ahlberg et al. 1992) but not map-based visualizations. In map-based visualizations these properties are generally linked to maps in traditional text format, which, as we argue, is not well-suited for a variety of visual tasks that are commonly performed on maps. To improve the architecture of map-based visualizations and facilitate the performance of visual tasks, we propose to integrate additional representations of ontological properties into digital maps.

A variety of representations have been empirically shown to be useful in a number of scientific map-based visualizations (e.g., Liao et al. 2006; Koua 2005; Guo et al. 2005; MacEachren et al. 2003). Additionally, multiple representations have been found to be effective for visualizing geospatial metadata (Ahonen-Rainio 2005). Many of these map-based visualizations are linked to self-organizing maps, parallel coordinate plots, multivariate representations, animations, and reorderable matrices. These representations facilitate both visual tasks with maps and visual exploration of patterns that are otherwise difficult to discover.

In the context of library and information science, the terms visualization and graphical representation are often used interchangeably, even though in other contexts, visualization implies only a set of graphical representations (i.e., visual artifacts that encode information in a graphical form and make effective use of human visual perception). In this paper, the interchangeable convention will be used, since the primarily concern is the discussion and selection of representations that facilitate visual tasks. In general, we consider a visualization to be more than just a set of representations; we are mostly interested in how people interact with representations.

This paper brings together research from a variety of fields, such as information visualization, geovisualization, human-computer interaction, and cognitive science, to address the issue of visualizing non-geospatial properties of documents with map-based visualizations of library collections. It introduces a conceptualization that merges isolated research directions and defines the new research foci in map-based visualization. We not only propose to use additional representations, but also examine how these representations can facilitate visual tasks. To demonstrate the benefits of this conceptualization, a prototypical interactive map-based visualization using GM has been developed.

The remainder of the paper is divided into three parts. In the next section, we review ontological properties of documents (books in particular), current practice of visualizating library collections using maps, and visual tasks that can be performed with these visualizations. In the following section, we describe possible representations of ontological properties of documents using GM. This part explains how individual properties can be encoded using different representations and which tasks can be supported by each representation. In the final section, we conclude the discussion and point to the future research suggested by our conceptualization.

2.0 Background Information

2.1 Ontological Properties of Documents and Other Objects

Libraries deal with many physical and abstract objects. These include documents or sets of documents such as bibliographic works or bibliographic families (Svenonius 2000; Smiraglia 2002; IFLA Study Group on the Functional Requirements for Bibliographic Records 1998), geographic entities, authors, subjects, and sometimes time periods. Documents and bibliographic works are generally the most significant of these objects; the rest are supplementary and traditionally have been used to support the organization of documents. Each of these objects with the exception of bibliographic works, have ontological properties that are captured in knowledge organization systems (KOS). Generally, properties of documents are described in item-level metadata records; properties of placenames are described in gazetteer records or placename authority records; properties of authors are described in authority records; and properties of subjects are described in classifications. Bibliographic works along with their expressions and manifestations are derived from entries in several KOS according to the Functional Requirements for Bibliographic Records (FRBR) that provide guidelines for determining the boundaries of a work (IFLA Study Group on the Functional Requirements for Bibliographic Records 1998). Gazetteers, authority records, and classification schemes are widely known in LIS as KOS (Hill et al. 2002; Zeng 2008). They are characterized by rich content and reside outside the collections of objects. They contribute controlled sets of labels for concepts, authors' names, definitions, and relationships to item-level metadata records. In this paper we extend the notion of KOS to item-level metadata standards such as Machine Readable Cataloguing (MARC). Although MARC is distinct from other content-containing KOS objects such as authority records and gazetteers, which we use to populate the fields in a particular MARC record, it has rich content and is used to organize knowledge similar to other KOS schemes. For this reason, we consider MARC as a special form of a KOS and will refer to all knowledge organization schemes as KOS throughout this paper.

The ontological properties of a book can be categorized as either physical or conceptual. Physical properties refer to the number of pages and the height of the book, descriptions of illustrative matter (presence of illustrations, maps, portraits), bibliographic notes, standard numbers, publication year, place of publication, edition information, and so on. Conceptual properties refer to titles, subjects (including geographic subjects), authors, publication types, call numbers, publication language, and relationships with other documents. According to the final report of the MARC Content Designation Utilization Project that studied commonly populated fields in MARC, the above-mentioned properties are the most frequently described in MARC records out of all published materials (including books, pamphlets, and printed sheet records) (Moen and Miksa 2007). Additionally, a subset of these properties is more or less common to all other metadata standards such as Dublin Core, International Standard Bibliographic Description (see crosswalks among ontological properties in various standards such as Network Development and MARC Standards Office, Library of Congress 2001; El-Sherbini 2001).

Metadata records reference various other KOS, each of which has its own unique structure. The Library of Congress *Classification* (LCC), for example, has a hierarchical structure composed of classes and subclasses. Each class has a name and a call number associated with it. The *Library of Congress Subject Headings* also has its own structure, which includes subjects, their synonyms, and variant spellings. Its subjects have various relationships to other subjects and are arranged into a hierarchy. Name authority records capture information about authors such as established standardized names, nicknames, pseudonyms, years of life, and references where the information about authors came from. Series and uniform title authority records contain information about any document that has been published under more than one title. The best way to describe geospatial properties is by using gazetteers (Hill 2006). Gazetteers are special kinds of KOS that describe properties of geospatial locations such as coordinates, placenames, variant names, and types of places. Once placenames are desribed in gazetteers, they can later be referenced in metadata records.

Ontological properties of documents have many usages. They are used for organizing and sharing knowledge, retrieving information, judging relevance, collocating bibliographic works, and comparing documents. KOS that record ontological properties set standard criteria according to which items can be compared and judged. Some conceptual properties are more important than physical ones for the retrieval and organization of knowledge, and users rarely name physical properties as important criteria in relevance judgements (Barry 1994; Savolainen and Kari 2006; Tombros et al. 2005; Xu and Chen 2006). These lesser-used physical properties, however, may provide information as to the probable structure and organization of key elements within a text. For instance, when readers pick up a book, they are afforded numerous aesthtic cues such as its size, age, condition, and number of pages, which can be indicators for effort involved in reading it, relevance, and previous usage rates, respectively (Dillon 2004; Reuter 2007). Often when one is searching for pictorial works, art albums, or atlases, the size of a book might become a valuable piece of information in determining its relevance. Determining and comparing book sizes in current map-based visualizations, however, is not a task which average library users generally know how to accomplish. But representing this information graphically can aid users in performing such tasks. For this reason, we emphasize that map-based visualizations should be KOS-fit. That is, they should represent more than just geospatial, temporal, and topical properties; rather they should represent other properties in order to provide users with the right information to perform quantitative and qualitative comparisons, as well as other visual tasks.

2.2 Representations

Map-based visualizations are hierarchical structures comprised of a variety of lower-level representations such as symbols, icons, text, images, diagrams, plots, tables, and even maps. Each type of lower-level representation has utility depending on the context in which it is used. For instance, in science education, diagrams can facilitate understanding of laws and problems (Chen 1996; Cheng 1992); in health sciences, symbols may help with symptom elicitation (Moriyama et al. 1994); and, in literacy education, symbols may provide cues to word meanings (Sheehy 2002; Fossett and Mirenda 2006).

Map-based visualizations often use base maps and other representations such as symbols, icons, timelines, labels, and subject representations. A base map is the ground layer upon which other layers of data are displayed. Examples of base maps are planimetric maps, topographic maps, and satellite views (Hill 2006). In map-based visualizations of KOS, base maps can be used to represent the structure of classifications (the LCC in particular) (Buchel 2006). For example, linguistic maps can be used as base maps for language and literature classes in the LCC, and historical maps for some historical classes.

Symbols and icons come in many different forms. Examples include markers, balloons (as used by GM), prisms (as used by Platial.com), graduated circles (as used by Flikr.com), 3D spatial histograms (Ancona et al. 2002), iconic stacks, differently-shaped coloured blocks (Ahonen-Rainio and Kraak 2005), and many others. When used in the context of maps, symbols and icons can make desired information perceptually salient. They can also make the relations between geospatial objects more readily apparent, which can facilitate higher-order knowledge-oriented activities such as problem solving, decision making, and hypothesis generation. When icons or symbols are viewed together, they should form a cohesive picture so that overall patterns in the multivariate information space can be discerned (Harris 1999; Dorling 1991). In other words, multiple icons should allow users to exploit the capacity to sense and discriminate the texture of a complex image (Pickett and Grinstein 1988). The purpose of icons at this level is to facilitate overall comparisons of objects and the identification of trends and unusual patterns in information (Harris 1999).

Most icons and symbols used in map-based visualizations of documents show counts, density, or simply serve as landmarks for objects.The purpose of icons such as Chernoff faces (Chemoff 1973), star glyphs, whisker plots (Ware 2004), and multidimensional icons (Spence 2007) is to encode rich ontological properties of documents. Such icons, when viewed together, can create interesting perceptual effects. For example, Chernoff faces may merge to form crowds, where one can easily compare faces to look for family resemblances or the mood of the crowd. Icons in the form of houses can be perceived as towns, allowing particular suburbs, estates, and streets to be identified (Harris 1999; Dorling 1991). The limitation of multidimensional icons, however, is their low dimensionality. Theoretically, the highest number of variables that Chernoff faces can encode is around 36 (Flury and Riedwyl 1981), but in reality, icons are hard to classify or associate when there are too many dimensions. Moreover, it has been found that not all facial features are equally distinguishable. De Soete and De Corte (1985) have demonstrated, for example, that only mouth curvature, eye size, and eyebrow density variables are the most salient. Ware (2004) recommends that star glyphs and whisker plots should use only a very small number of orientations (possibly 3), in order to facilitate rapid classification and association of glyphs. In addition, Chernoff faces may arouse emotions among subjects and may lead to confusing interpretations (Ahonen-Rainio 2005).

Labels should be an integral part of any graphical representation. According to Tufte (2001, 180), "Viewers need help that words can provide. It is nearly always helpful ... to label outliers and interesting points." Additionally, label sizes and colours may facilitate the grouping of semantically-related labels, and may guide one's attention and increase the possibility of learning something from the graphical representation (Götzelmann et al. 2005a). For users that are not familiar with the geography of the visualized region, map labels may aid in the exploration of relative locations of cities using cardinal directions, and relations between cities in terms of distances which are missing in alphabetical arrangements in the LCC.

Tufte (2001) also recommends placing labels directly on the graphic itself without the use of legends, because "words and pictures belong together, genuinely together" (Tufte 1990, 116). Legends may prevent users from seeing each part and its label together. Other detailed aesthetic and legibility requirements for positioning geographic labels on maps are described in Imhof (1975). To summarize, geographic labels should be legible, be easily associated with features they describe, reflect the hierarchy of features by the use of different font sizes, and not be densely clustered, evenly dispersed, or overlap other map contents. The hierarchy of cities is traditionally derived from population sizes. Cities with larger populations have larger labels, while cities with lower populations have smaller labels.

Timelines have attracted much interest from researchers in the areas of digital libraries and geovisualization over the past two decades. While digital library researchers have investigated how to intergrate temporal gazetteers (also known as time directories) with map-based visualizations (Zerneke et al. 2006), geovisualization researchers have been working on the graphical representation of time in geographic information systems (Edsall 2001; Li and Kraak 2008; Peuquet 2002). The proposed representations can encode time in a linear, cyclical, combination of the two, or even static manner. In the age of print maps, for example, time was typically encoded using labels (Vasiliev 1997). For the visualization of documents, both linear time encodings and text labels are suitable, although the former is preferred. Together, as the linked maps and other displays change, KOS and graphical timelines allow zooming and scrolling while the linked maps and other displays change to reflect the state of information at a moment in time.

Subject and theme representations are also part of map-based visualizations. Information visualization researchers have used a variety of techniques for visualizing subjects. For example, subjects can be represented as a Kohonen Map (Kohonen 1995), Tree Map (Shneiderman 1992), Topic Map (Weerdt et al. 2007; Le Grand and Soto 2002), Flickr-Style Tag Cloud (Bausch and Bumgardner 2006), Cartogram (Keim et al. 2003), Fisheye View (Furnas 1986), Hyperbolic Tree (Lamping et al. 1995), and others. These representations have the aesthetic appearance and proper balance between visualization models and underlying data, as strongly connected nodes appear close to each other, and weakly connected nodes appear far apart (Chen 2004). The scalability issue, however, is one of the serious drawbacks of these representations. As the size and density of data increases, most of these representations do not scale well.

Other subject representations such as spatialization techniques have been introduced by cartographers. An example of a spatialization is ThemeScape (Wise et al. 1995). Spatializations use spatial metaphors such as distance-similarity, direction, scale, arrangement, regionalization, and landscape (Montello et al. 2003; Fabrikant et al. 2004). These metaphors are built on analogy with the real world and allow users to intuitively explore the space (Börner et al. 2003). Spatializations typically employ dimension reduction techniques and layout algorithms to project similarity in subjects onto distance, such that semantically-similar subjects are placed closer to one another (Börner et al. 2003).

Map-based visualizations cannot be complete without interactive legends. With print maps legends typically contain a heading and labels, legend boxes, or symbols to depict numerical values or nominal classes, class units, and explanatory text (Sieber et al. 2005). Static legends are often criticized for impeding visual tasks, because users have "to dart back and forth between textual material and the graphic" (Tufte 2001, 180). On the other hand, digital maps, due to their interactive nature, do not have this shortcoming, as they are mobile and dynamic. Digital legends are no longer used solely for decoding purposes. They can be used for modifying the appearance of the map and the classification of the data, for performing information retrieval (querying the thematic layers), and for filtering the results (data extraction and data suppression) (Sieber et al. 2005). As to the design of legends, surprisingly it is not easy to find any clear guidelines or recommendations for their design, except for the framework for the design of self-adaptable legends for digital atlases by Sieber and colleagues (2005).

Ontological properties of documents in metadata are traditionally linked to maps either as tabulated text or tables (see an example in Figure 2). When information elements are encoded in a table, exact details are more easily extracted because users know which field to attend to (Newton 1994), and juxtaposed cells can highlight differences (Winn 1987). Tables are usually more effective than graphical encodings for small datasets (Tufte 2001). As tables get larger, however, search is hindered by a large number of distractors with multiple features. For this reason, if a precise match is found in a large table, the user has no way to know whether or not it is the best match (Spence 2007).

Tables and tabulated text make it easier for users to "visually parse" the metadata fields, but they do not facilitate reasoning about relative locations of documents and relationships among them. For instance, tabulated text may describe concrete physical and relative locations of documents, but without encoding the locations graphically, inferences can be difficult to make (Brown and Dowling 1998). Veerasamy (1997) suggests another problem with such textual encodings: metadata records can only be viewed individually

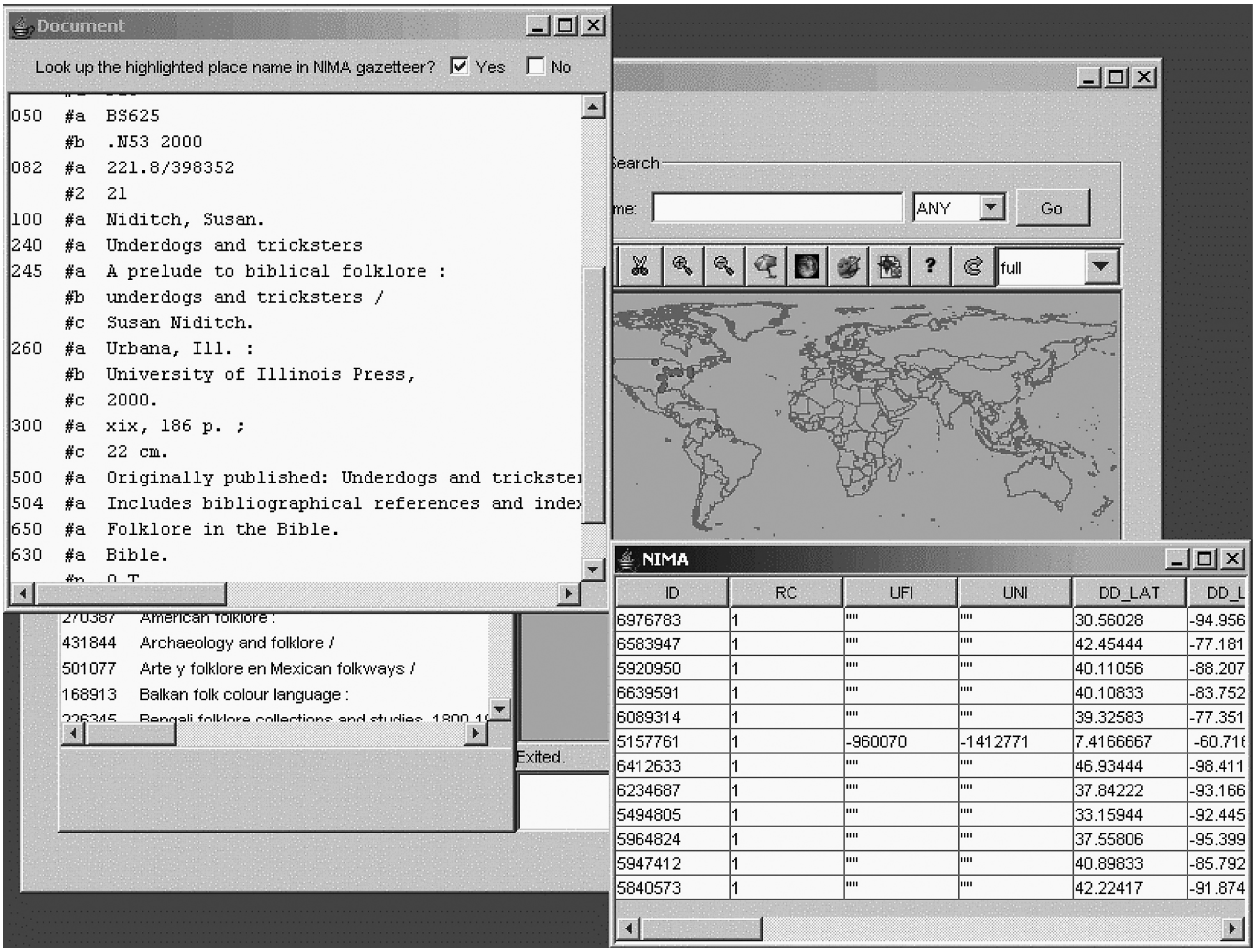


Figure 2. A map-based visualization with links to metadata and gazetteer records which are represented either as tabulated text or a table (Electronic Cultural Atlas Initiative 2004).

instead of as a set, which does not allow for distinguishing and grouping documents based on their commonalities and idiosyncrasies. Additionally, tables often list only key ontological properties (e.g., title, subject, the year of publication, and call number for documents), thus preventing users from comparing physical properties of documents. As a result of these shortcomings, users often engage in shallow processing of the information space while extracting elements from tables (Newton 1994; Vonk and Noordman 1990).

To facilitate the comprehension of ontological properties, additional graphical representations can be used. Modern digital maps such as GM and Google Earth allow symbols and icons to be linked to other representations (e.g., other maps, images, plots, charts, self-organizing maps, diagrams, timelines, and legends). In general, representations can be linked by either one of three linking types: replacing, overlaying, and repeating (Wilhelm 2008). Depending on the context in which they are used, different linking techniques can either enhance or impede visual exploration of representations (Wilhelm 2008; Plaisant 2005; Sedig et al. 2005). With the first type of linking, replacing, old information is typically lost and gets replaced by new information. The shortcoming of such linking is that relevant previous states must be held in one's memory if they are to be integrated with new knowledge (Wilhelm 2008). Replacing is used in cartographic space-time animations, which are composed of a series of image frames, where each image appears later in time than the previous (Bétrancourt and Tversky 2000). The downside of these animations is that they may cause change-blindness, since users have to keep track of multiple events occuring simultaneously that require them to split their visual attention (Lowe 1999; 2003; Fabrikant et al. 2008). For example, Rensink and colleagues (1997) have demonstrated that users have great difficulty noticing even large abrupt changes between two successive scenes in an animation, when blank images are shown between scenes. This can be overcome, however, when changes in representations are caused by user interaction (Sedig et al. 2005; Wilhelm 2008).

In the second type of linking, overlaying, the new information is placed on top of the original representation. Typically, map overlays are opaque, conform to the scale of the original map, and are georegistered to the geographic location on the main map. Figure 3 shows a map where overlaying is used. In this example the map is overlayed on an aerial image. Such overlays help users to compare the accuracy of representations, or to develop a composite representation of a location, combining both the original representation and the overlay. Identical scales in linked maps and georegistration provide a common framework for comparing the linked layers. Overlaying facilitates the comparison of linked representations, since the differences between the two are so readily apparent. Examples of where overlaying has been used are in map and aerial imagery collections (e.g., as described in Forbes and Janée 2007, and in the David Rumsey Map Collection at http://www.davidrumsey.com/).

In the third type of linking, repetition, the displays are juxtaposed and different representations of the same data are available. This is a preferred type of linking for many map-based visualizations. Such visualizations place juxtaposed representations on the screen, and their changes are synchronized (see an example in Figure 4.a). The advantage of this type of linking is that it provides a rather comprehensive picture of the data; the user has a complete overview of every representation and can clearly observe the impact of changes in one representation on another. Problems with this type of linking include lack of screen space

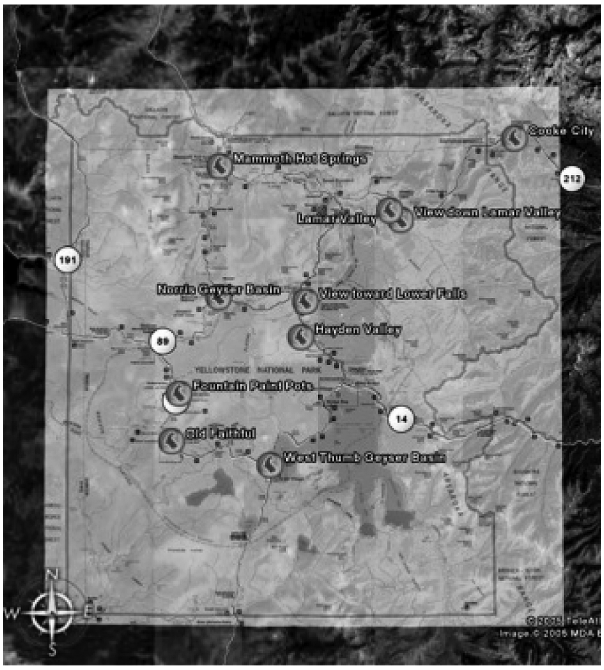
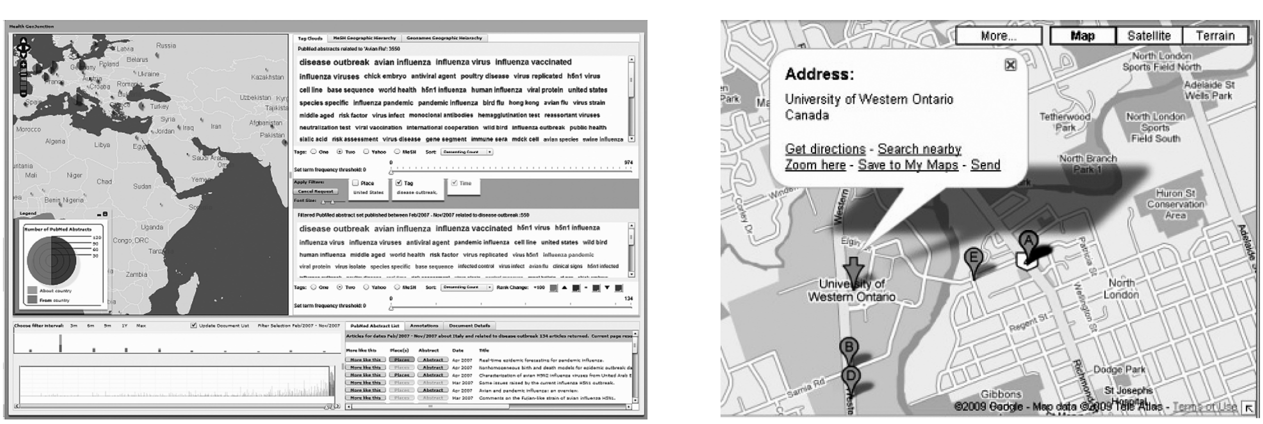


Figure 3. A map overlay. Such overlays help users to compare the accuracy of representations, or to develop a composite representation of a location, combining both the original representation and the overlay.
Source: http://naturalsciences.org/microsites/education/treks/yellowstone/images/places/index.html.

and an often unintuitive interface. Users frequently need help to understand the interface, especially when using an unfamiliar application (Andrienko N. et al. 2002). To overcome the complexity of juxtaposed representations, Plaisant (2005) proposes to use the metaphor of sticky notes overlaid on top of the linked views. This is meant to help locate the main widgets,



a) b)

Figure 4. Examples of linking by repetition: a) repetition in juxtaposed layers and frames; b) repetition via information windows in GM.

demonstrate their manipulation, and explain the resulting actions.

GM use a different solution for overcoming complexity. They allow linking additional graphical representations via information windows which can be opened or closed on request (as shown in Figure 4b). This is similar to linking by repetition, but not in its pure form, since the linked view does not remain on the screen all the time. The drawback of such linking is that it prevents comparisons between multiple linked views.

2.3 Visual Tasks

Traditionally, textual representations of ontological properties are used to support information retrieval tasks, such as finding, identifying, selecting entities, and acquiring or obtaining access to the entities (Svenonius 2000). However, retrieval tasks are not enough. Users should also be able to engage in visual tasks. These include scanning the visual scene; detecting patterns, shapes, graphs, and edges; discriminating labels, text, color, motion, properties of objects and surfaces (relative size, magnitude), and relationships; discriminating 2-D shapes from 3-D objects; grouping similar objects and regions; focusing on a spotlight; recognizing landmarks (MacEachren 1995); relating distances (semantic similarity or geospatial relevance) and values; identifying clusters (Koua 2005); and reading and comparing values (Edsall 2001). The most basic visual tasks performed while reading maps are identifying, locating, comparing, and associating (Andrienko G. et al. 2005). While identifying, users search for thematic characteristics of objects; while locating, users search for positions of objects in time and space; while comparing, users establish qualitative and quantitative relationships; while associating, users relate attribute values or spatial patterns (Andrienko G. et al. 2005). While information retrieval tasks can be delegated to a computer, visual tasks can only be performed by users. The efficacy of supporting visual tasks depends not only on the designer's skills to represent information, but also on individual user's cognitive, spatial, and visual abilities.

Visual tasks can be further divided into elementary and synoptic exploratory tasks (Andrienko N. and Andrienko G. 2006). Elementary tasks refer to identifying individual elements of the reference set. For example, finding the value of an attribute corresponding to a certain specified reference such as how many books a library has about a particular city or how the books in a collection are distributed geographically. Synoptic tasks involve the identification of the whole reference set or its subsets. In such tasks, sets of references are considered in the entirety of their charateristics (which include behaviours and relations between reference sets or between behaviours). Here, behaviours refer to distributions, variations, and trends. An example of a synoptic task would be determining the variation of document counts over a whole county or country. Among synoptic tasks, the most challenging are tasks of finding connections between phenomena (e.g., cause-and-effect relationships or finding the principles of the internal organization, functioning, or development) (Andrienko N. and Andrienko G. 2006).

The possibility of map-based visualizations to support such a variety of visual tasks increases their utility in comparison to text-based library catalogs. Whereas well-designed map-based visualizations can reveal implicit relationships among library collections, traditional textual representations can only help with identifying and locating items, associating documents with hard-coded relationships, and navigating between geographic places by means of broader and narrower relationships. Maps, on the other hand, allow users to see locations in time and space, learn about nearby cities, distances between cities, directions, changes over time, and so on. Despite these additional advantages, maps do not facilitate visual tasks for comparing and associating documents. In order to design visualizations that support visual tasks, maps must be augmented with additional graphical representations that allow users to identify, compare, associate, and locate documents visually. This will allow users to perform not only elementary tasks of finding items, but also synoptic tasks which allow them to see trends and patterns in publication and acquisition data. This is important for scalable map-based visualizations with hundreds and thousands of metadata records linked to each city.

3.0 Prototype Map-Based Visulization System

In this section we describe our selected testbed document collection, our prototype system (Section 3.1), and our design of graphical representations that encode ontological properties of documents (Section 3.2).

3.1 Selected Document Collection

For the purposes of this research, we have selected an experimental collection of documents. Our collection is comprised of 349 records from the Library of Congress catalog. We used the LCC structure to guide our selection of records to visualize, since geographical and chronological arrangements in the LCC are framed in accordance with the needs of each subject field. The LCC has been "praised for the freedom allowed in each schedule for development according to its subject field's own intrinsic structure" (Angell 1964). This intrinsic structure can facilitate the development of interesting map-based visualizations for each subject field, which can hardly be derived from simpler structures. Each individual geographic arrangement along with other related classes in the LCC can be characterized as a mini-ontology of a certain theme. Some of these mini-ontologies include georeferences (i.e., geographic names), temporal references, and other elements. These elements can be grouped together in map-based visualizations. An example of a geographic mini-ontology in the LCC is the "Local History and Description" classes. Ranges of geographical references about the "Local History and Description" classes are dispersed throughout the schedules D and F of the LCC. They list various countries, cities, and other placenames, to which resources about the "Local History and Description" of those places are linked. Some of these ranges are presented in Table 1.

Country	LCC schedules
Afghanistan	DS374-375
Austria	DB101-879
Canada	F1001-1145.2
China	DS781-796
Finland	DL1170-1180
India	DS483-486
Italy	DG600-975
Japan	DS894-897
Korea	DS936-937
Portugal	DP702-802
Russia	DK511-651
Spain	DP285-402
Switzerland	DQ301-857
Ukraine	DK508
USA	F10-975

Table 1. LCC Classes About Local History and Description (Library of Congress, n.d.)

This list is far from complete, as there are many more ranges like these. Some of these ranges have more georeferences than others. Among georeferences within the specified ranges, one can identify geographic concepts of various scales (ranging from countries, regions, and cities to small geographic locations).

Documents in the "Local History and Description" mini-ontologies share many common features. First, they have similarities in bibliographic descriptions, as they have similar subjects and forms/genres, are written in the same languages, are published in close locations, have similar call numbers, and are placed on the shelves close to each other. Second, many of these items contain illustrative matter (illustrations, maps, portraits, and so on), because they relate the history of geographic places. Third, some of the items have large formats, which can be of great interest to users (e.g., oversized pictorial works). Fourth, in terms of forms, cataloguers place travel guides, encyclopedias, directories, guidebooks, pictorial works, tour guides, and other similar items in these sections. Fifth, they have no temporal features (i.e., items within these ranges are not arranged according to time periods).

All records in our testbed collection have call numbers starting from DK508. Class DK508 is an example of the mini-ontology described above. All records within this mini-ontology have the same subject: "Local History and Description" of Ukraine. All documents within this class are about Ukraine, including regions and cities. For example, subclasses DK508.922-DK508.939 are assigned to documents about Kiev, DK508.95.L86 to documents about L'viv, and DK508.95.O33 to documents about Odessa. There are a few exceptions where a subclass includes documents about more than one geographic location, but these are rare. An item about L'viv (title: Lwów, przewodnik / Ryszard Chanas, Janusz Czerwiński), for instance, has a call number starting with DK508.924, which belongs to the range of call numbers assigned to Kiev. This item was eliminated from the subset during the uploading in the research prototype. Among downloaded records there are the entire subsets of records for 32 Ukrainian cities (among which the largest cities are Kiev, Odessa, L'viv, and Kharkiv). The numbers of MARC records linked to each city subclass in DK508 range from a few records to almost a hundred.

From each MARC record only the key ontological properties were selected. They are physical descriptions (illustrations, maps, height of the book, number of pages in the book, year of publication, place of publication—all of these attributes are recorded in

MARC 300), languages (field 041), types of publication (serial or monograph, field 006), bibliographic notes (MARC 504), subjects (fields 650 and 651), titles (field 245), call numbers (field 050), and acquisition numbers (MARC 010). Other properties, such as standard numbers, geographic area codes, uniform titles, variant titles, edition statements, publication and distribution information, series statements, note fields, were considered beyond the scope of this research, because of their low occurrence rates in the testbed collection. All records were downloaded into a MySQL database.

Georeferences in the LCC are indirect; that is, for describing locations on a map, each placename requires a mathematical representation in the form of geospatial coordinates. As was mentioned in Section 2.1, the best practice to describe geospatial properties is by using gazetteers. In our prototype, we used a locally-built gazetteer which includes footprint information (latitude and longitude), names (main and variant), and feature types represented by population sizes.

As records were downloaded, the raw data was very carefully inspected and potential problems with the representation of certain fields were identified. For example, book pagination usually has mixed notation (e.g., 81 p., [8] p. of plates; xviii, 204 p., [1] folded leaf of plates). Such notation is difficult to summarize computationally. For now, only values recorded as Arabic numerals were used in the visualization. More experimental work is required to decide what to do with the remaining notations.

Another problem with the testbed dataset is missing values, typical of many datasets. In book pagination, missing values transpire more often than in the other fields. Instead of pagination, 300 |a, a field may carry information about the number of volumes. The following are the examples of such problematic records, with the field italicized:

|**a** *v. <1>* : |**b** ill. ; |**c** 17 cm. (DK508.935 |**b** .V68 2002)
|**a** *v. <1-6>* : |**b** ill., maps ; |**c** 17 cm. (DK508.935 |**b** .K68 2002)
|**a** *2 v.* : |**b** ill. (some col.) ; |**c** 20 cm. (DK508.934.K67 A3 1998)
|**a** *v. <1-2 >* : |**c** 20 cm. (DK508.933 .K96 1997)

In this paper, we do not address the issue of missing values, as a total of six problematic records with missing values are simply omitted from the visualization. Our omission is justified, as missing values in visualizations are not uncommon. For instance, Ahonen-Rainio (2005) reported that she had missing values in geospatial metadata and had to assign some special codes to the problematic items to make them visible on the display. This helped the users know that the records existed, but that their fields were missing. Some researchers suggest that allowing users to know which values are missing is as important as allowing them to know which ones exist (Tang et al. 2003).

The prototype collection does not include serial records, since their ontological properties are significantly different from the ontological properties of books. To test the visualization design of serials, a larger dataset of serials would be required. The number of serials in our collection was only 2; therefore, they were excluded. We also did not include any documents about regions and countries, since their visualizations might require different base maps.

Lastly, when selecting records for the testbed collection, we noticed that some cities included in the LCC no longer exist. Examples include the following names of the cities or locations of ancient fortresses: Chufut Kale (Çufut Qale), Eski-Kermen, Feodosiia Kaffa, and Chersonese. Some of these placenames are names of the cities that existed on the territories of modern cities. For example, the modern town of Feodosiia occupies the sites of ancient cities of Theodosia (which existed between the 6th century B.C. and the 4th century A.D.) and Kaffa (which existed between the 13^{th} to 18^{th} century). It was named Feodosiya only in 1802 (Feodosiya 2007). This is an example where cities temporarily overlap.

To visualize ontological properties of documents, our prototype system uses GM as the visualization platform. All representations and interactions described in this paper are developed with the GM API, PHP, Ajax, and Fusion Charts (http://www.fusioncharts.com/). Although ideas and principles outlined in this paper could be implemented using other geographic information systems, we chose GM because they already have many base layers as well as built-in linking and customization capabilities.

3.2 Representations of Ontological Properties

In this section we review additional representations, how they can be linked, what visual tasks they can support, and what ontological properties they can encode. In particular, subsection 3.2.1 describes how to add missing labels for placenames, why they are necessary, and what visual tasks these labels can support in map-based visualizations. Subsection 3.2.2 discusses the design and role of symbols. The design of timelines and the visual tasks they support is presented in subsection 3.2.3. Subsections 3.2.4-3.2.5 explain the possible design and application of digital legends. Finally, the concluding subsection 3.2.6 explores the utility of additional graphical representations in map-based visualizations.

3.2.1 Placename Labels

The most important ontological property of documents for map-based visualizations is georeferences. Because GM has a built-in gazetteer and has labels for the majority of placenames, many developers of map-based visualizations approach the design of maps with the assumption that displaying labels for georeferences is not necessary. Although this is true, GM is composed of 18 map layers with various scales and therefore different sets of labels. Due to variations in scales, some layers have names of the cities, while others do not. On some layers, cities are encoded as dots, and on others as small polygons. While designing map-based visualizations, developers choose the most appropriate layers. For instance, layers suitable for the representation of cities of various countries and regions are usually layers 5-10 with GM. On these layers, countries and city locations can be viewed at the proper scale, and users do not have to pan maps in order to view all cities on one map. But many of these layers lack all placename labels, because the size of the location is not big enough to be represented on all map scales. Thus, when mapping cities of Ukraine on layer #7 (which is the most suitable for the visualization of cities in Ukraine), only 17 out of 32 cities that have to be represented have labels. Cities with populations smaller than 500,000 people are not shown on this map layer.

Thus, in order to represent cities with populations smaller than 500,000 people, we could use the following typology:

Cities of less than 10,000
Cities of 10,000 to 100,000
Cities of 100,000 to 200,000
Cities of 200,000 to 500,000.

The names of larger cities can be encoded using a larger font than the names of smaller cities. See Figure 5 for demonstration of how this approach works on GM, where added placename labels are circled in red.

To provide users with clues about non-existing cities, the labels for such cities may either be shown in a different colour (e.g., red or grey), or visualized with the help of a timeline. In the case of "overlapping"

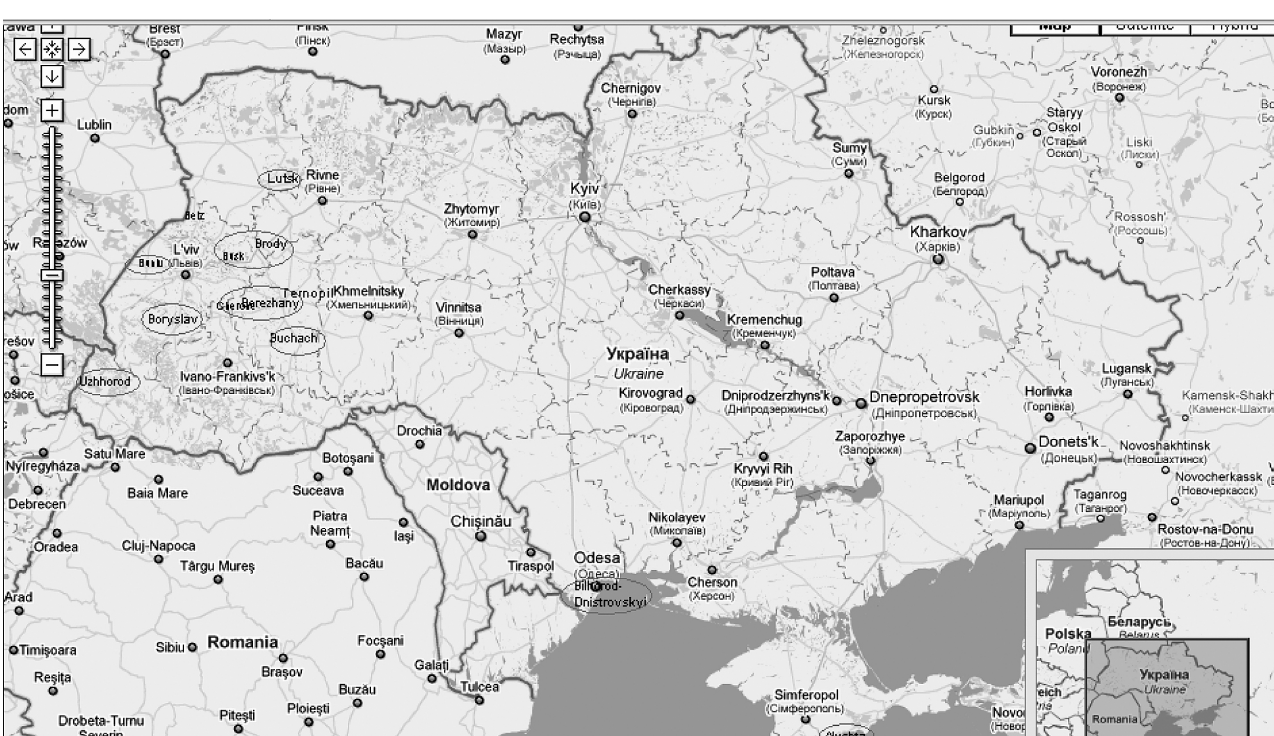
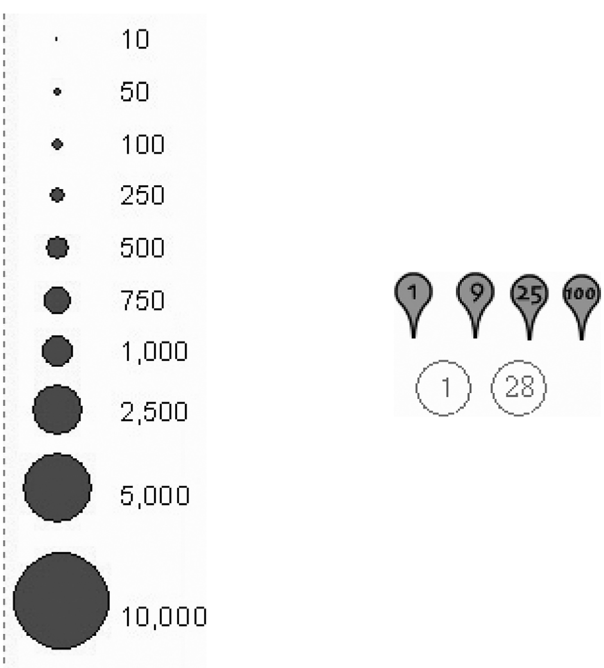


Figure 5. A map with added placename labels.

cities, developers could also use sunflower symbols for locations of the cities (e.g., ⅄ + ✳). A sunflower symbol is a dot with short, radiating lines called petals. The number of overlapping cities can be represented by the number of petals (Cleveland and McGill 1984; Harris 2007).

3.2.2 *Symbols in Map-Based Visualizations*

To represent documents, designers traditionally use symbols and icons. Symbols and icons may be neutral (e.g., simple markers), or reveal either the quantity (density) or the semantics of linked objects (i.e., their ontological properties). Examples of quantitative symbols and icons include graduated symbols, or icons with numbers denoting quantities or densities. Quantitative symbols may show counts, averages, sums, mins, maxs, medians, or orders.



a) b)

Figure 6. Examples of quantitative symbols: a) graduated symbols; b) icons with numbers.

At first, it might appear that quantitative values are not relevant to ontological properties of documents that are mainly categorical. Categorical properties, however, can be transformed into counts by mapping the nominal values to numbers (i.e., assigning order and spacing to the nominal values) by means of statistical operations (such as count, average, sum, min, max, median), sorting, or filtering (in order to display only interesting ranges of data) (Rosario et al. 2003). For example, in map-based visualizations, graduated symbols and icons with counts may show quantities of items retrieved for each city (see Figure 7). The advantage of markers with actual counts over graduated symbols is that actual counts can support quantitative comparisons in datasets where items are not equally distributed. For example, one of the methods used in the design of graduated symbols is the division by quartiles, where the first 25% of datapoints are shown with the smallest symbol, the next 25% with a larger symbol, and so on (Slocum et al. 2005). For instance, let us assume that in our collection items 1-10 constitute the first quartile. This means that no matter whether the city has 1 or 2 or 3 or 7 matching items, all of them will be shown as one symbol. This example demonstrates that graduated symbols do not allow users to differentiate between the quantities within the quartile ranges, whereas actual counts can easily do so.

Quantitative symbols support elementary visual tasks (e.g, locating cities and identifying quantities of items associated with each city), including quantitative comparisons of collections. For example, on the map in Figure 7 we can identify the cities with the highest densities (Kyyiv-89, L'viv-87, and Odessa-53), the ones with the lowest densities (Lugansk, Rivne, Boryslav, Brody, Cherche, Buchach, Nikolaev have only one item each), and the ones without any collections. Moreoever, symbols become visible after a brief exposure which lasts typically from 30 to 300 milliseconds and is accomplished during preattentive processing (Ware 2004; Spence 2007). The advantage of using symbols that exploit such processing is that it does not require conscious attention and does not overload working memory.

Text labels combined with quantitative symbols may assist users in quantitave comparisons of collections. Counts of objects make it easy to determine which collections are larger and which are smaller, and scaled placename labels facilitate making comparisons among the cities. Logically, the density of collections is higher for well-known, large cities, and lower for small- or medium-sized cities. Any deviations from this may indicate errors, gaps, or other unusual situations.

Multidimensional or semantically-rich icons can encode different ontological properties of documents. Examples of such icons include Chernoff faces (Chernoff 1973), Star Icons (Ward 1994), and fourfold/twofold displays (Friendly 2000; Stryker et al. 2008). Unfortunately, we were not able to encode ontological properties of datasets for each city. The

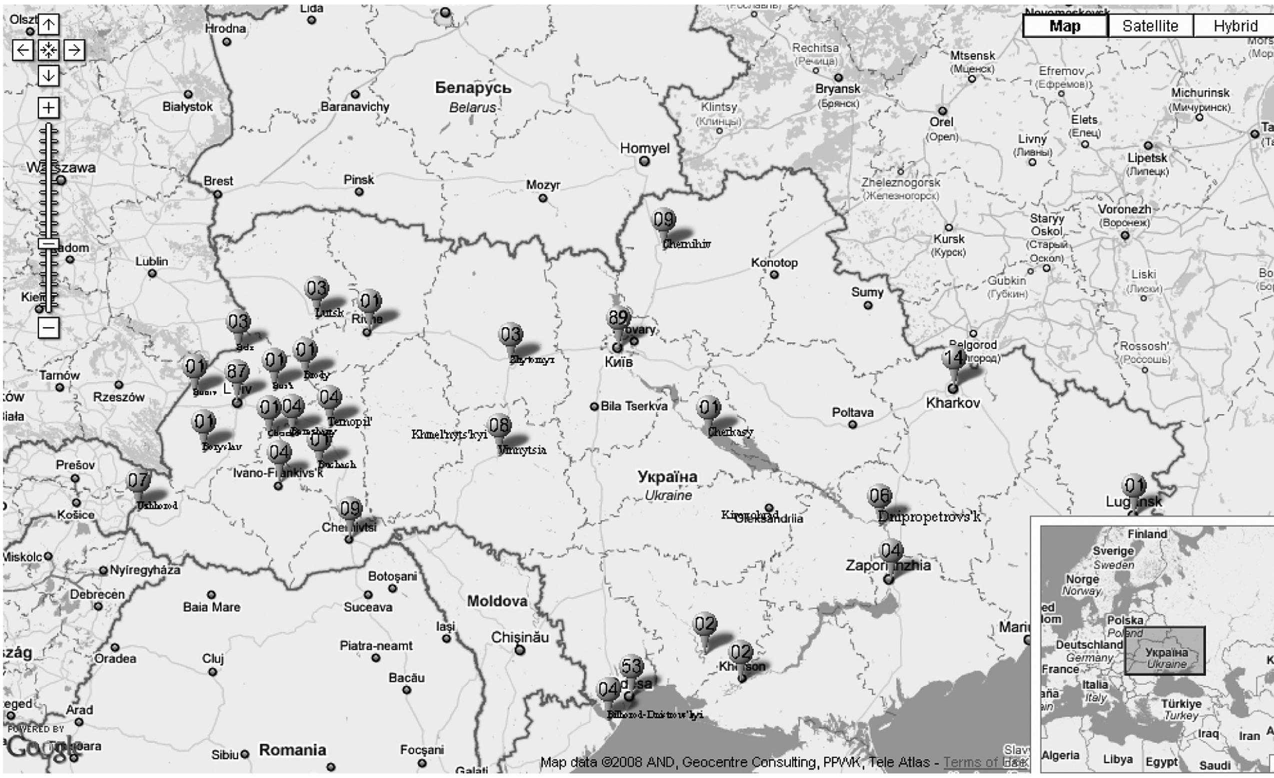


Figure 7. Quantitative symbols on Google Maps.

problem with radar plots, Chernoff faces, and fourfold/twofold displays is that they are only effective for the representation of datasets with high-volume counts. They can hardly be useful for cities with scarce collections such as the one we had selected. We were able to generate multidimensional icons only for 4 cities, since other cities (28 out of 32 cities) have collections less than 20 items.

3.2.3 Timelines

Timelines can be used for displaying and tracking events, objects, and activities (Kapler and Wright 2004). Many map-based visualizations of documents typically display only a single dimension of time, most often temporal aboutness (i.e., the time period of the document as in ECAI TimeMap). KOS, however, have many other temporal aspects that can be displayed and tracked on timelines, such as publication years, acquisition years, and authors' lifelines. The provision of such timelines can aid users in performing elementary visual tasks involving questions such as

- What was the first book acquired by this collection?
- Which books were acquired in 2007?
- Books about what cities were written by the authors who lived between 1920-1930?
- How recent are the library collections?
- What books were written by contemporary authors?
- When were most of the books published?

To visualize several temporal aspects, designers in information visualization often use hierarchical timelines (André et al. 2007). Hierarchical timelines can include several linked timesliders, each of which has a time scale and control handles that allow setting a selected range on the time scale (see Figure 8 below). To faciliate elementary visual tasks, timelines often include histograms or line plots showing counts of items associated with each time period. Histograms allow users to identify years or time periods with the highest density of activities, events or objects. To ease the quantitative comparisons between different timelines, all timesliders should use identical time scales.

Besides elementary visual tasks, hierarchical timelines can allow users to perform an array of visual synoptic tasks, which are not possible with simpler visualizations. Due to the specifics of our prototype collection, however, we did not have enough data for the visualization of all temporal aspects (i.e., temporal

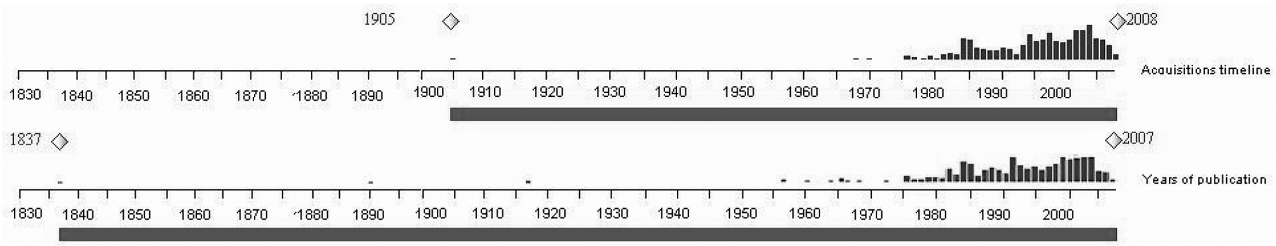


Figure 8. Hierarchical timelines. These timelines represent acquisition and publication years.

aboutness, acquisition timelines, publication timeline, and authors' lifelines). We were able to visualize only acquisition and publication years. But even these two timelines are sufficient to explain and demonstrate the richness of visual synoptic tasks that can be supported when multiple timelines are used and/or linked together. Synoptic tasks are not simply about identifying objects and reading their values; they are more about the comparison of behaviours (Andrienko N. and Andrienko G. 2006). For example, synoptic tasks with acquisition and publication timelines allow users to assess acquisition and publication trends over certain time periods. More specifically, users can answer to the following types of questions:

- Which books, published during the Soviet Period, were acquired after 1991?
- When were the books acquired in 2004 published?
- What is the time-lapse between publication and acquisition of a set of books?

3.2.4 Legends

Digital legends are commonly linked to maps either in separate frames or layers. They are linked through repetition and are synchronized with the map so that any change in the legend causes changes in the map, and vice versa. The appearance of digital legends is not always static; rather, they can be self-adaptable (Sieber et al. 2005). Such legends adapt their shape and content to various conditions (e.g., to scaling and repositioning of various layer set-ups). Additionally, adaptive legends can adjust their content to include only symbol categories that actually appear on the map. In this case, the legend may update dynamically according to navigation.

Legends in current map-based visualizations of documents sometimes afford switching between different base maps (e.g., a linguistic map, a physical map, or a religious map), filtering by country or city, and selecting keywords, subjects, or topics. The results of the analysis of ontological properties in our prototype collection suggest that such legends have limited capabilities for interaction with ontological properties of documents. For example, besides subjects and keywords, documents can have information about forms/genres, languages, formats, illustrations, maps, and bibliographic notes.

In our prototype, the legend (see Figure 9) displays the most frequently occurring ontological properties: forms recorded in form subdivisions (e.g., biographies, dictionaries and encyclopaedias, gazetteers, guidebooks, pictorial works, tours); frequently used subjects; and physical attributes (e.g., illustrations, maps, portraits, bibliographic notes). The legend may also include other properties such as categories of placenames (e.g., extinct and modern, or population sizes) and languages.

Representating ontological properties in the legend has many advantages. Without the legend users do not know what is linked to the map-based visualization. Properties allow users to preview hidden structures in the database. Moreover, users can see what properties (subjects, types, and physical properties) appear more than others in the metadata records linked to the LCC DK508 class. In other words, the legend serves as a semantic lense for these records.

While visualizing entire library collections, in which many other similar classes (i.e., mini-ontologies about the Local History and Description) can be found, the legends may be adapted dynamically to include only categories extracted from the metadata MARC records associated with each mini-ontology in the LCC. It is highly possible that mini-ontologies could be associated with slightly different forms, subjects, and physical properties of documents. Some of the advantages of using dynamic legends are: 1) they demonstrate the association between the LCC classes and geographic regions that they describe; 2) they generalize the structure of each class in the LCC; and, 3) they allow users to observe differences in ontological properties among the LCC classes. For example, in MARC, there are as many as 484 discreet language codes (Library of Congress

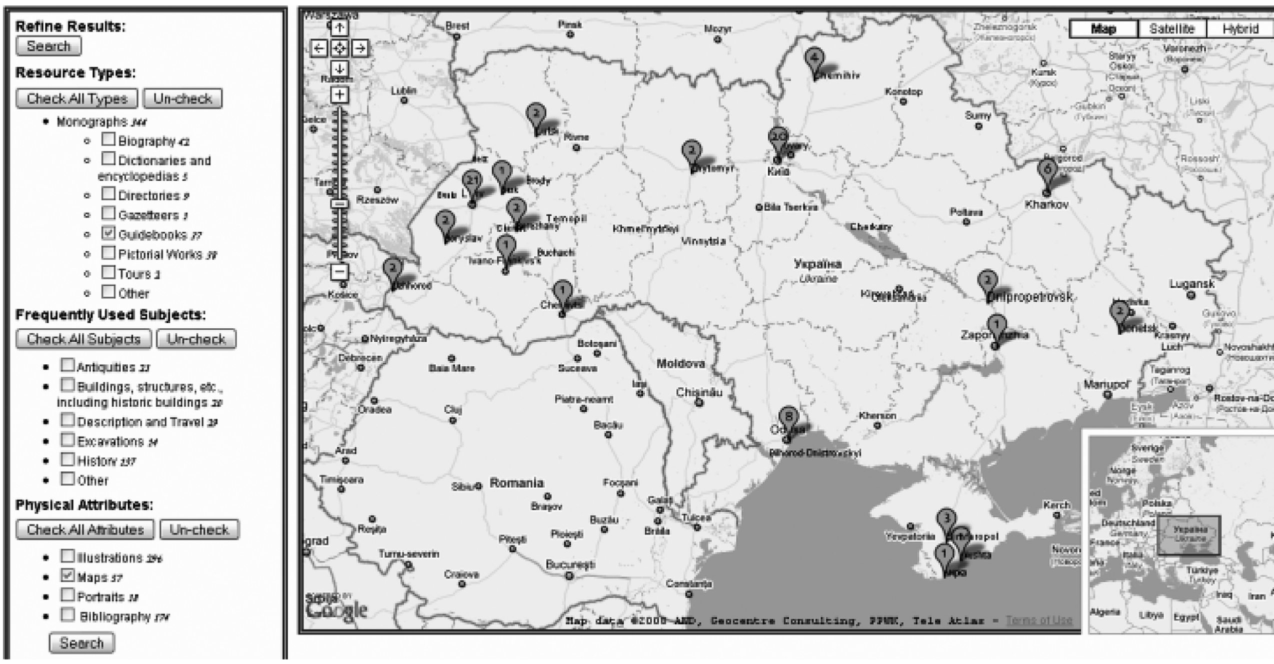


Figure 9. A map with a legend. This snapshot shows what the legend retrieves when users select Guidebooks and Maps categories in the legend.

Network Development and MARC Standards Office 2007). It is unrealistic to include all of them in a static legend next to the map. Moreover, not all of these languages are relevant to each geospatial region. Self-adapting legends allow each geospatial region to be associated with a unique set of languages relevant only to that region.

In addition, self-adapting legends enable users to pose complex dynamic queries using a mouse, knowing neither the details of the underlying database schema nor the details of first-order logic. They allow users to select on any legend element and, thereby, set the query parameters to the values of the properties of that element. Furthermore, a self-adapting legend, as conceptulized in our prototype, serves as a control mechanism that allows users to regulate how many objects to retrieve on the map. If too many items are retrieved, the query parameters on the legend can be adjusted to make the number of results smaller. Conversely, when no items or just a few items are retrieved, more relevant criteria can be checked on the legend.

Together with quantitative symbols, the legend facilitates the visual exploration of the retrieved results, which is supported by elementary and synoptic tasks. Together, they allow users to immediately preview how items are distributed and decide whether the number of results is satisfactory or not. They allow users to discover relationships between topical subjects (e.g., history or archeology) and geographical subjects, as well as forms of documents and their geospatial distributions.

Our legend is different from traditional legends by the nature of the categories (or ontological properties) it represents. While traditional maps have non-overlapping categories (one object has one attribute), on our map categories can overlap in metadata records. One metadata record can have one or more physical attributes, one or more subjects, and one or more forms/genres. This means that theoretically, regardless of how many categories are specified in the legend, only one item on the map can be retrieved. For example, Figure 9 demonstrates what the legend retrieves if we select guidebooks (the count for which is 37) and *maps* (the count for which is 57). Logically, users would expect the total number of retrieved results to be the sum of these two, which is 94. But adding up the numbers on quantitative symbols gives us the surprising result of 82, an explanation for which is that 12 guidebooks have maps. A limitation of our legend is that it allows adding parameters with only the Boolean operator "AND", with no option to search and retrieve items excluding some of these criteria. Graphical represention of Boolean queries is still an open research question in information visualization.

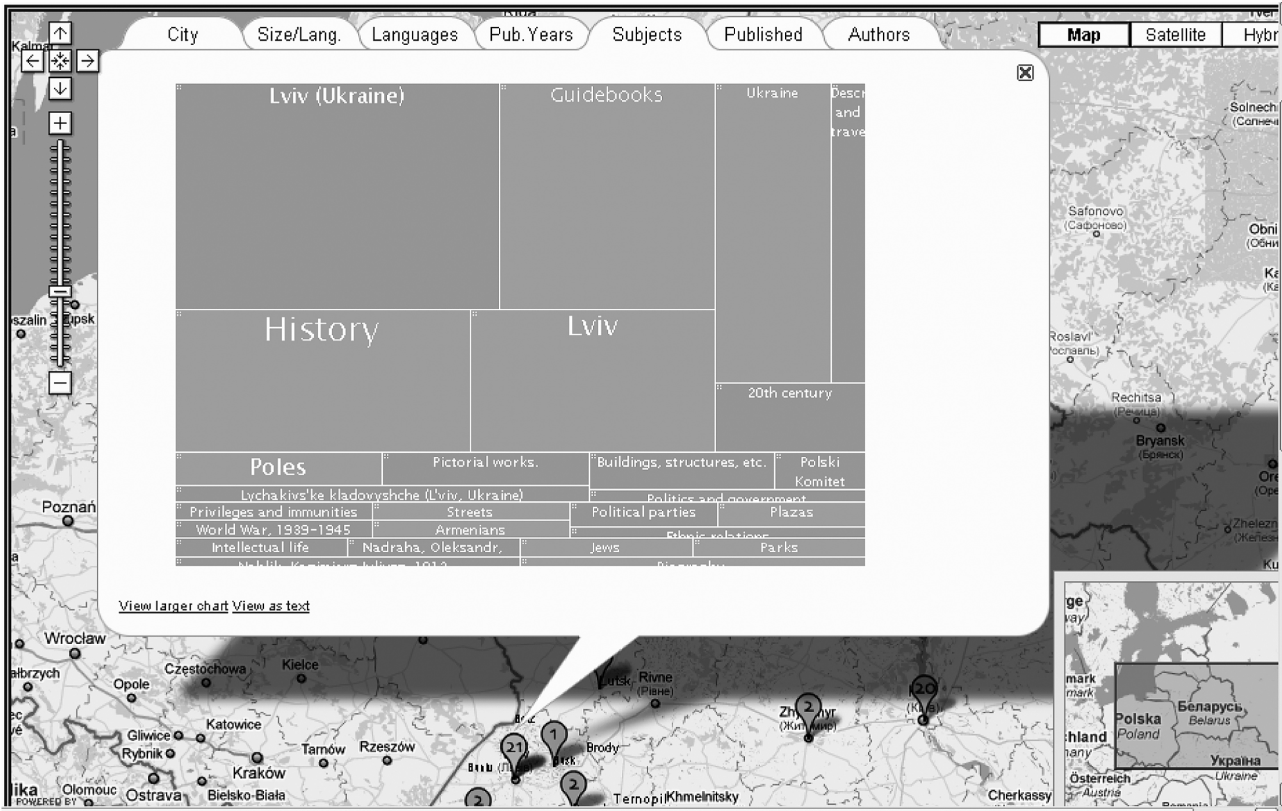


Figure 10. Filtering Kohonen Maps by linking to information windows in GM. This figure shows how graphs, spatializations or Kohonen Maps can be filtered by placename and linked to information windows for each city.

3.2.5 Legends and graphical representations

Legends can include both textual and graphical representations. Sometimes graphical representations can completely replace text as spatializations, generalized graphs (e.g., trees and networks), or self-organizing maps such as Kohonen maps or tag clouds. These graphical representations in legends have both advantages and disadvantages. Their advantages include providing users with an additional view using a single visualization that captures a complete picture of the data space. The scalability issue in generalized graphs, spatializations, and tag clouds is a major disadvantage for the visualization of subjects from library catalogues (Chen 2004 and 2005). Even in our dataset comprised of 343 records, the number of distinct subjects is 574. Presenting all of these subjects next to a geographic map would not be effective. First, there would not be enough room for displaying a large network of subjects. Second, a single global tag cloud as a primary means of partitioning is imprecise and has low recall (Hayes et al. 2006). Third, even if they are displayed, many low-volume subjects do not have labels, which renders them almost useless to users.

To reduce the complexity of graphs, spatializations, and tag clouds, researchers in information visualization and geovisualization have proposed several solutions. Information visualization researchers suggest to divide (or filter) representations into smaller components, to prune (Chen 2004), or to thematically cluster them (Lohmann et al. 2009; Hayes et al. 2006). Thematically-clustered tag clouds can assist with discriminating secondary information to further refine and confirm the knowledge produced by the clustering. Furthemore, clustered clouds establish topic-based relationships between tags that were not observable when considering global tag clouds alone (Hayes et al. 2006). Geovisualization researchers use zooming and panning into relevant areas of spatializations or graphs (Hubmann-Haidvogel et al. 2009), or filter representations by time (Stryker et al. 2008). Although useful, such geovisualization techniques have a few shortcomings. For example, in zooming and panning, visualizations focus on certain areas of graphs, spatializations, or tag clouds; as a result, the focus does not include all relevant information; additionally, time-filtered tag clouds may not show thematic relationships among topics.

Pruning and dividing the graphs can be useful for the visualization of subjects in legends. For example, subjects in MARC records include not only topical subjects (e.g, history and archeology), but also geographic subjects. Showing geographic subjects in the legend might not be necessary, since each geographic subject will be shown on the map anyway. Another type of pruning and filtering can be achieved through a different type of linking. As explained above, legends are usually linked to maps through repetition and are synchronized with maps. In addition, GM provide other linking possibilities; namely, additional graphical representations can be linked via information windows for each city. Each representation linked to a city this way will automatically have a smaller number of subjects (or any other ontological properties) to display, simply because it will be a representation of one city, not the entire mini-ontology (see filtered Kohonen map in Figure 10). The downside of such linking is that it will be difficult to compare representations linked to different cities. But it is unclear whether comparisons among collections for different cities are relevant to the map-based visualization of ontological properties of documents. Such comparisons are crucial for the analysis of geospatial phenomena, but probably not as crucial for comparisons of ontological properties.

3.2.6 Additional Representations

Kohonen maps, graphs, spatializations, and tag clouds are not necessarily the best representations of ontological properties of documents. Other graphical representations may be better at providing insight into other document properties at a glance (e.g., book formats, languages, authors, etc.). The strength of such an approach is the possibility of quickly imparting the various aspects of the collection structure from KOS to users.

The first additional representation which we suggest to add to map-based visualizations is the scatterplot. The scatterplot is one of the most powerful tools for data analysis (Cleveland and McGill 1984) due to its simplicity, familiarity, and visual clarity (Elmqvist et al. 2008). Scatterplots are distant relatives of maps, since like maps they map the information space onto a two-dimensional space (Friendly and Denis 2000). Much like maps, scatterplots show the amount of association between variables, the dependence of variables, cluster of points, outliers, among other things (Elmqvist et al. 2008).

Scatterplots assign data dimensions to graphical axes and render data cases as points in the Cartesian space defined by the axes (Elmqvist et al. 2008). To map categorical values onto scatterplots, designers often transform nominal values by changing them to numbers (i.e., assigning order and spacing to the nominal values) (Rosario et al. 2003). Such transformation procedures allow scatterplots to be used extensively in information visualization (e.g., Ahlberg et al. 1992; Ahlberg and Shneiderman 1994a; Ahlberg and Shneiderman 1994b; Kang and Shneiderman 2000; Nowell 1997). The number of dimensions that a single scatterplot can reliably represent, however, is considerably less than many realistic datasets. To rectify this, scatterplot visualizations often give some control to users to dynamically switch between the visualized dimensions (Nowell 1997). For example, in Nowell (1997), possible combinations of axes are: document type and relevance, author and publication year, and index terms and relevance.

Among ontological properties in our collection, however, we have two numerical variables which naturally possess the inherent order, spacing, and distance necessary for being mapped onto scatterplot axes. These properties are: the number of pages of a book and its height. Both of these properties suggest the shape or format of the book. Following Tufte's recommendation "If the nature of the data suggests the shape of the graphic, follow that suggestion" (Tufte 2001, 190), we plotted book width on the x-axis, and height on the y-axis (see an example in Figure 11). Our scatterplot positions books in a metaphorical space that provides insight about individual book sizes, similarities, and dissimilarities among books in terms of their formats. The representation of book sizes on the scatterplot helps users direct their attention to items of interest, to oversized books, or thick books, just as users would do this if books were placed on a bookshelf. It also allows users to draw conclusions about the average book size and the outliers.

Using symbols, small icons, or letters to encode their datapoints, scatterplots can expand their analytical affordances. Namely, they can assist users with perceiving the point cloud of a particular category as a unit as if the other points were not there. There is a great variety of techniques available to represent datapoint categories. They are letters, filled or unfilled simple geometric shapes, simple shapes such as the circle and regular polygons, sunflowers (Cleveland and McGill 1984), points of the same subset connected by lines (Becker et al. 1987), radial line sym-

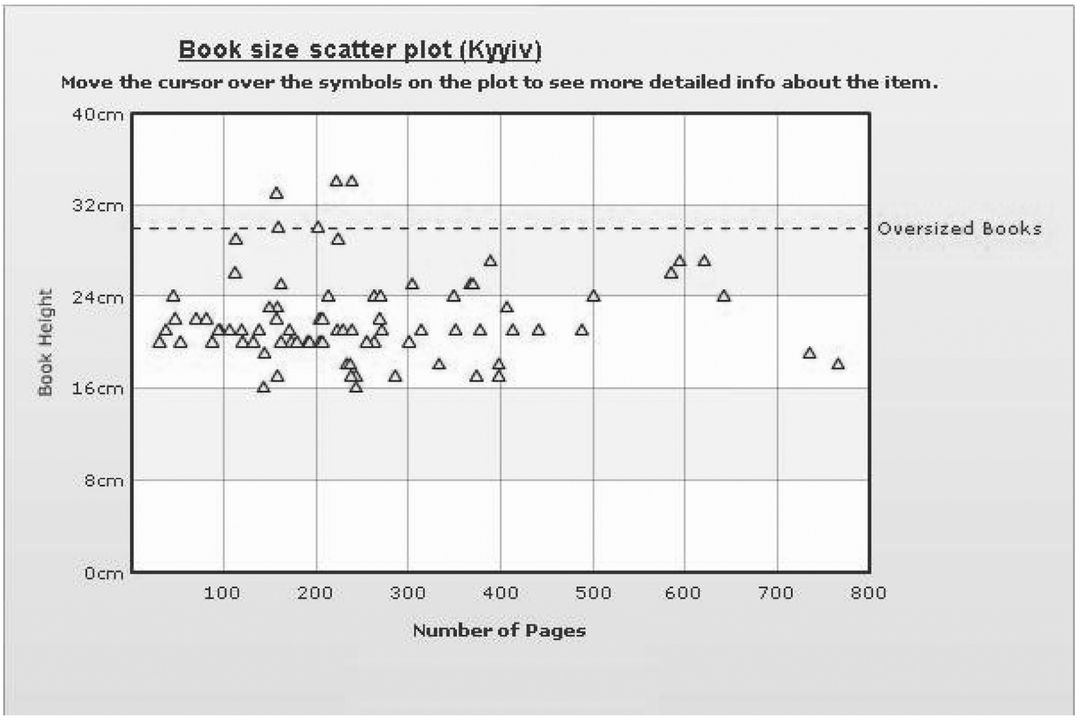


Figure 11. A book size scatterplot. This figure shows sizes of 89 books. From this scatterplot users may conclude that approximately two-thirds of all books have fewer than 300 pages; 3 books are taller than 32 cm; and 6 books have more than 500 pages. The height of most books is around 20 cm.

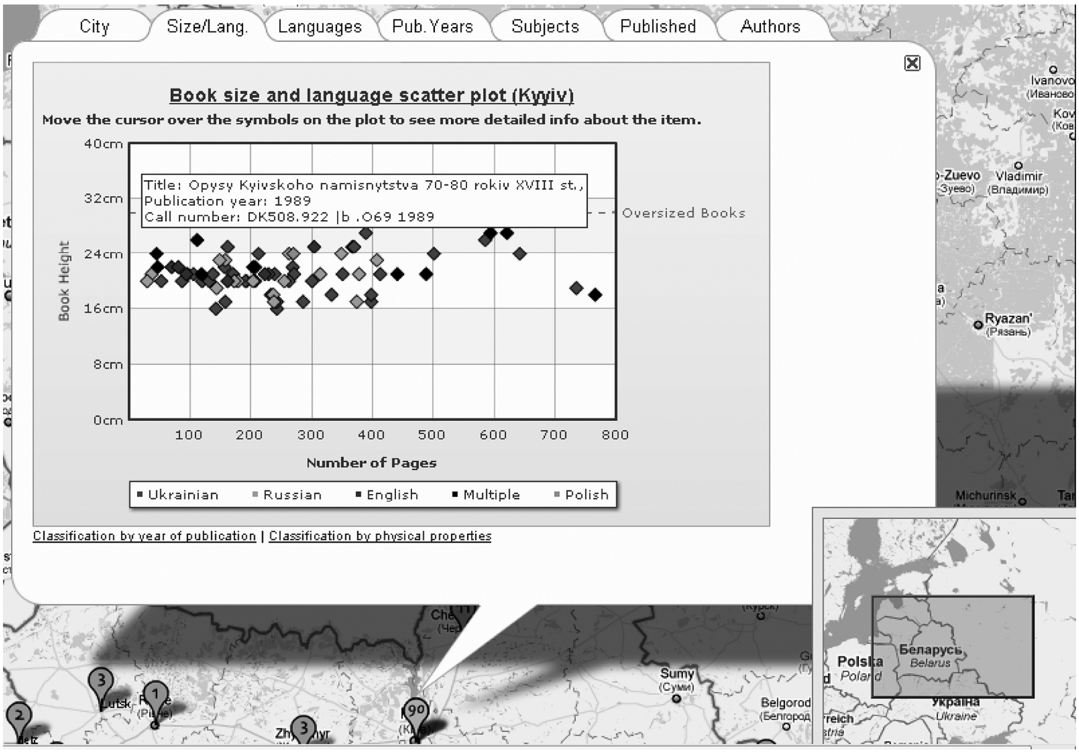


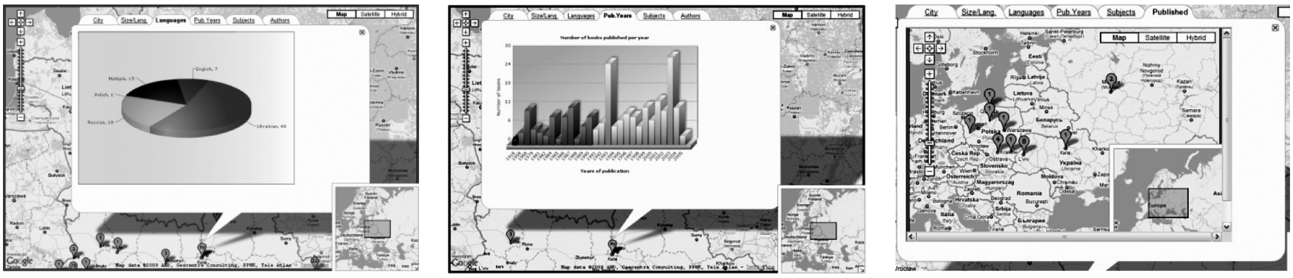
Figure 12. A colour-coded scatter plot. Colour-coded geometric shapes on this scatterplot show languages of documents about Kiev. There are a few items in Russian, English, and Multiple languages.

bols, the plus sign, or the asterisk (Tremmel 1995). For example, Figure 12 shows how languages can be encoded as geometric shapes of different colours. Colour-coded shapes will help users see hidden properties of items and will help them relate similar objects, determine relevant objects, and perceptually estimate the distribution of relevant items without any extra cognitive effort. Coloured lines between symbols can be used for encoding bibliographic relationships among various editions or other manifestations of bibliographic works, if items have different dimensions. In addition, symbols can have built-in textual descriptions of items to scaffold users in the interpretation of graphics (e.g., titles, years of documents, or even thumbnails) and may have graphically represented links to other expressions of the same bibliographic work (e.g., videos or CDs). If more than one document has the same dimensions, the symbol size can be used to encode multiple documents.

Other supplementary representations that can be useful are pie-charts, histograms, and embedded geographic maps. These representations may be used for representing languages, local timelines, and even geographical maps (see Figure 13). Each of these representations reveals the structure of separate ontological properties and supports different visual tasks. Thus, the pie chart, unlike the scatterplot, shows the proportion of languages in the whole set (see Figure 13a). This representation supports quantitative comparisons of documents based on languages.

Histograms may show local timelines (see Figure 13b). While global timelines for publication and acquisition years record purchasing and publishing events in the entire DK508 mini-ontology, dynamically-generated local timelines, linked to information windows for each city, can expose chronological gaps, temporal extents, and currentness of datasets for each city. For example, in our dataset, histograms for each location vary by the start date. For some places, the Library of Congress has resources published in the 19^{th} century (e.g., Odessa's collection was published starting from 1837), while for others the first collected items were published more than a century later (e.g., Kharkiv's collection starts from 1986). Furthermore, each location may have its own set of historical periods (or events), which can be added to KOS for visualization purposes, with historical time periods being colour-coded. For example, the vertical bars in the Kiev histogram show such time periods as "before the October revolution of 1917" (shown in black), "the Soviet times" (shown in red), and "after 1991" (shown in yellow). Colour-coded histograms not only can show statistics and assist with quantitative comparisons, but also may emphasize trends in publishing (e.g., during the Soviet times, the publication rates about local history were low, and they became higher after 1991 when Ukraine became independent).

Geographic maps can be used to provide links to themselves. This is possible since documents have two major geographical properties. The first is the geographic aboutness, which is about subject or content; the second is the geographic place of publication. Typically, map-based visualizations provide equal access to the geographic aboutness and the place of publication, and users are allowed to switch between these two maps. However, when not viewed together, users are limited in the visual tasks they can perform. The map of publication places allows users to determine the spatial distribution of publishers and possibly the volume of documents purchased from each publisher, and the map of geographic aboutness allows users to see the distribution and density of geographical subjects. But when publication places are coupled with geographic aboutness, they can tell users much more. In particular, they can tell users information such as where books about a particular city



a) b) c)

Figure 13. Additional representations: a) a pie chart showing languages of documents about L'viv; b) a histogram of publication rates by year (city Kiev); c) a map with places where documents about L'viv (Travelguides and items with maps) were published.

were published. For example, the map of publication places in Figure 13c shows users that half of travel guides and items with maps about L'viv were published outside Ukraine, with 9 items published in Poland. Such multiplicity and crosslinking of representations, we argue, extends the utility of map-based visualizations, supporting a more diverse set exploratory and visual tasks to be performed on document collections.

Besides these representations, map-based visualizations can have links to graphical representations of bibliographic works and citations. Bibliographic works can be represented as tree-like graphs that sort and arrange bibliographic record sets according to the FRBR model (one such visualization is described in Leazer and Furner 1999). In our prototype dataset, however, the majority of bibliographic works have a single manifestation. This is not surprising, according to a study by Bennett et al. (2003), 78% of works in library catalogs have a single manifestation. The works that have the most manifestations can probably be found in other parts of a collection such as literature, psychology, and religion. Self-adapting legends in map-based visualizations can possibly facilitate visualization of FRBR model in the contexts where it is relevant and omitting it in the others.

In addition, map-based visualizations can be linked to citation maps represented either on a Self-Organizing Map, or with a means of factor analysis, multidimensional scaling, or eigenvalue decomposition techniques (Börner et al. 2003). Such geospatially-contextualized citation maps can help reveal the role of culture, identity and collaboration in the research communities (Chen et al. 2008).

It is important to note that it is difficult to add disciplines to the visualization at this stage. The problem is that disciplines are at the very high level in the LCC. The documents presented in this paper are all about history and they are linked to the lower-level classes in the LCC.

Last but not least, it is assumed that all the representations discussed in this paper are interactive. Static, non-interactive, representations are limited in aiding human cognition (Spence 2007). Interactions should at least enable users to probe and retrieve different elements of the map-based visualizations. Otherwise, much of the semantic and relational properties of these visualizations will remain hidden and latent (Sedig and Sumner 2006).

4.0 Conclusions and Future Work

Throughout this paper, we have drawn attention to two research issues concerning map-based visualizations. The first research issue is the need to extend map-based visualizations to include representations that encode ontological properties of documents. The proposed representations in our prototype are not necessarily the most usable, effective, or efficient representations of ontological properties. Testing is underway to determine effectiveness of these representations. The second research issue, tightly coupled with the first, is the need to take into consideration visual tasks that can be supported by visual representations. Currently, when designers create map-based visualizations of documents, they are primarily concerned with mapping individual items on digital maps. They are less concerned with how people understand these items, and what visual tasks these mappings can facilitate. To overcome these limitations, we need to reconceptualize map-based visualizations in the context of representations and visual tasks. This clearly points to the need for design frameworks for map-based visualizations, classifying representations and their associated tasks. Developing such taxonomic design frameworks should be an important objective of map-based visualization research.

Ultimately, map-based visualizations need to address a fundamental question about how computers can amplify human cognitive and epistemic abilities. In the context of geospatial references, visualizing concepts and time periods is an insufficient goal. Map-based visualizations must also aid users in performing epistemic activities, such as making sense of linked library collections and generating hypotheses about collections.

The visualization of ontological properties of documents may take many forms and this paper has only provided an introduction to such visualizations. Due to the limitations of our prototype dataset, we were not able to demonstrate how map-based visualizations can be augmented with the representations of citations, bibliographic works, disciplines, and some other document properties. It is hoped that further research will construct and test more sophisticated visualizations that incorporate additional ontological properties and support other visual tasks and epistemic activities. This will improve the design of map-based visualizations and facilitate rich interactions with geospatial data.

References

Ahlberg, Christopher, and Ben Shneiderman. 1994a. Visual information seeking using the filmfinder. *In CHI '94: Conference companion on Human factors in computing systems.* New York, NY, USA: ACM Press, pp. 433–34.

Ahlberg, Christopher, and Ben Shneiderman. 1994b. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. *Proceedings ACM CHI'94: Human Factors in Computer Systems*, pp. 313-17.

Ahlberg, Christopher, Christopher Williamson, and Ben Shneiderman. 1992. Dynamic Queries for Information Exploration: An Implementation and Evaluation. *Proceedings ACM CHI'92: Human Factors in Computer Systems,* pp. 619-26.

Ahonen-Rainio, Paula. 2005. Visualization of Geospatial Metadata for Selecting Geographic Datasets. Unpublished dissertation for the degree of Doctor of Science in Technology, Helsinki University of Technology , Espoo, Finland.

Ahonen-Rainio, Paula, and Menno-Jan Kraak. 2005. Deciding on fitness for use: evaluating the utility of sample maps as an element of geospatial metadata. *Cartography and geographic information science* 32n2: 101.

Ancona, Dan, Michael Freeston, Terrence Smith, and Sarah Fabricant. 2002. Visual explorations for the Alexandria Digital Earth Prototype. In *Visual interfaces for digital libraries*, edited by K. Börner and C. Chen. Berlin: Springer-Verlag, pp. 199-213.

André, Paul, Max L. Wilson, Alistar Russell, Daniel A. Smith, Alistar Owens, and m.c. schraefel. 2007. Continuum: designing timelines for hierarchies, relationships and scale. *UIST2007 (ACM Symposium on User Interface Software and Technology).* Newport, Rhode Island.

Andrienko, Gennady, and Natalia Andrienko. 2007. Visual data exploration: Tools, principles, and problems. In *Classics from IJGIS. Twenty years of the International journal of geographical information science and systems*, edited by Peter F. Fisher. Boca Raton, Fla.: Taylor & Francis, pp. 475-79.

Andrienko, Gennady, Natalia Andrienko, Jason Dykes, Mark Gahegan, David Mountain, Penny Noy, Jonathan C. Roberts, Peter Rodgers, and Martin Theus. 2005. Creating instruments for ideation: software approaches to geovisualization. In *Exploring geovisualization*, edited by Jason Dykes, Alan M. MacEachren and Menno-Jan Kraak. New-York: Elsevier, Ltd., pp. 103-26.

Andrienko, Natalia, and Gennady Andrienko. 2006. *Exploratory analysis of spatial and temporal data: A systematic approach.* Berlin: Springer.

Andrienko, Natalia, Gennady Andrienko, and Peter Gatalsky. 2005. Impact of data and task characteristics on design of spatio-temporal data visualization tools. In *Exploring geovisualization*, edited by Alan M. MacEachren, Menno-Jan Kraak and Jason Dykes. New York: Elsevier, Ltd, pp. 201-22.

Andrienko, Natalia, Gennady Andrienko, Hans Voss, Fatima Bernardo, Joana Hipolito, and Ursula Kretchmer. 2002. Testing the usability of interactive maps in CommonGIS. *Cartographicand geographic information science* 29: 325-42.

Angell, Richard. 1964. Development of Class K at the Library of Congress. *Law library journal* 57: 353-54.

Archaeological Computing Laboratory - University of Sydney. 2004. *TimeMap: time-based interactive mapping*. http://www.timemap.net/ (accessed March 20, 2009).

Barry, Carol L. 1994. User-Defined Relevance Criteria: An Exploratory Study. *Journal of the American Society for Information Science* 45: 149-59.

Bausch, Paul, and Jim Bumgardner. 2006. Make a Flickr-Style Tag Cloud. In *Flickr hacks*. O'Reilly Press.

Becker, Richard A., William S. Cleveland, and Allan R. Wilks. 1987. Dynamic Graphics for Data Analysis. *Statistical science* 2: 355-83.

Bennett, Rick, Brian F. Lavoie, and Edward T. O'-Neill. *The concept of a work in WorldCat an application of FRBR.* . Dublin, Ohio: OCLC Office of Research, 2003.

Bertin, Jacques. 1983. *Semiology of graphics: Diagrams, networks, maps.* Translated by William J. Berg. Madison, WI: University of Wisconsin Press.

Bétrancourt, Mireille, and Barbara Tversky. 2000. Effect of computer animation on users' performance: A review. *Le Travail Humain* 63: 311-29.

Börner, Katy, Chaomei Chen, and Kevin Boyack. 2003. Visualizing Knowledge Domains. *Annual Review of Information Science and Technology*. Medford, N.J. Information Today, pp. 179-55.

Brown, Andrew, and Paul Dowling. 1998. *Doing research/reading research: a mode of interrogation for education.* 1. Falmer Press: Bristol, PA, 1998.

Buchel, Olha. 2006. Uncovering Hidden Clues about Geographic Visualization in LCC. Edited by Gerhard Budin and Christian Swertz. *Proceedings of the 9th International Society for Knowledge Organization Conference, Vienna, Austria*. Würzburg: Ergon-Verlag, pp. 77-84.

Buckland, Michael, Aitao Chen, Frederic C. Gey, Ray R. Larson, Ruth Mostern, and Vivien Petras. 2007. Geographic search: catalogs, gazetteers, and maps. *College and research libraries* 68: 376-387.

Buckland, Michael, Frederic Gey, and Ray Larson. 2002. *Support for the learner: what, where, when, and who. proposal narrative.* 2004. http://ecai.org/imls2002/ (accessed June 7, 2009).

Chemoff, Herman. 1973. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association* 68: 361-68.

Chen, Chaomei. 2004. *Information visualization: beyond the horizon.* 2nd ed. London: Springer.

Chen, Chaomei. 2005. Top 10 Unsolved Information Visualization Problems. *IEEE computer graphics and applications* 25: 12-16.

Chen, Peter C.-H. 1996. Problem solving and learning with diagrammatic representations in physics. In *Forms of representation: an interdisciplinary theme for cognitive science*, edited by Donald Peterson. Intellect, Ltd., pp. 47-66.

Cheng, Peter C.-H. 1992. Diagrammatic reasoning scientific discovery: Modelling Galileo's Kinematic Diagrams. In *AAAI technical report on reasoning with diagrammatic representations (Report mNo. SS-92-02)*. Menlo Park: CA: AAAI, pp. 33-38.

Cleveland, William S., and Robert McGill. 1984. The Many Faces of a Scatterplot. *Journal of the American Statistical Association* 79n388: 807-22.

De Soete, Geen, and Wilfried De Corte. 1985. On the perceptual salience of features of Chernoff faces for representing multivariate data. *Applied psychological measurement* 9: 275-80.

Dillon, Andrew. 2004. *designing usable electronic text ergonomic aspects of human information usage.* Boca Raton, FL: CRC Press LLC.

Dorling, Danny. 1991. The Visualization of spatial structure. PhD Thesis, Department of Geography, University of Newcastle upon Tyne.

Edsall, Robert M. 2001. Interacting with space and time: desigining dynamic geovisualization environments. Thesis, College of Earth and Mineral Sciences, The Pennsylvania State University.

Electronic Cultural Atlas Initiative. 2004. Going places in the catalog: improved geographic access. Final report, University of California, Berkeley.

Elmqvist, Niklas, Pierre Dragicevic, and Jean-Daniel Fekete. 2008. Rolling the dice: multidimensional visual exploration using scatterplot matrix navigation. *IEEE transactions on visualization and computer graphics*: 1141-48.

Fabrikant, Sarah I., Daniel R. Montello, Marco Ruocco, and Richard S. Middleton. 2004. The distance similarity metaphor in network display spatializations. *Cartography and geographic information science* 31: 237-52.

Fabrikant, Sarah I., Rebich Stacy, Daniel R. Montello, Gennady Andrienko, and Natalia Andrienko. 2008. A visual analytics approach to evaluate inference affordance from animated map displays. *GIScience 2008 Pre-Conference workshop on Geospatial Visual Analytics.* PArk City, UT.

Feodosiya. 2007. *The Columbia electronic encyclopedia.* Columbia University Press. http://encyclopedia2.thefreedictionary.com/Feodosiia (accessed July 1, 2009).

Flury, Bernhard, and Hans Riedwyl. 1981. Graphical Representation of Multivariate Data by Means of Asymmetrical Faces. *Journal of the American Statistical Association* 76n376: 757-65.

Forbes, Angus, and Greg Janée. 2007. Visually Browsing Georeferenced Digital Libraries. *Geoinformatics 2007 Conference.* San Diego, California.

Fossett, Brenda, and Pat Mirenda. 2006. Sight word reading in children with developmental disabilities: A comparison of paired associate and picture-to-text matching instruction. *Research in developmental disabilities* n27: 411-29.

Fox, Edward A., N. Dwight Barnette, Clifford A. Shaffer, Lenwood Heath, William C. Wake, Lucy T. Nowell, Jan Lee, Deborah Hix, and H. Rex Hartson. Progress in interactive learning with a digital library in computer science. *Proceedings ED-MEDIA 95, World Conference on Educational Multimedia and Hypermedia.* Graz, Austria, 1995, pp. 7-12.

Fox, Edward A., Deborah Hix, Lucy T. Nowell, Dennis J. Brueni, William C. Wake, Lenwood S. Heath, and Durgesh Rao. Users, user interfaces, and objects: Envision, a digital library. *Journal of the American Society for Information Science* 44: 480-91.

Friendly, Michael. 2000. *Visualizing categorical data.* Cary, NC: SAS Publishing.

Friendly, Michael, and Daniel Denis. 2005. The early origins and development of the scatterplot. *Journal of the history of the behavioral sciences* 41: 103–30x.

Furnas, George W. 1986. Generalized fisheye views. *Human Factors in Computing Systems CHI '86 Conference Proceedings*, pp. 16-23.

Götzelmann, Timo, Kamran Ali, Knut Hartmann, and Thomas Strothotte. 2005. Adaptive Labeling for Illustrations. *13th Pacific Conf. on Computer Graphics and Applications,* pp. 64-66.

Götzelmann, Timo, Kamran Ali, Knut Hartmann, and Thomas Strthotte. 2005. Form Follows Function: Aesthetic Interactive Labels. Edited by L. Neumann, M. Sbert, B. Gooch and W. Purgathofer. *Computational aesthetics in graphics, visualization and imaging*, pp. 193-200.

Guo, Diansheng, Mark Gahegan, Alan M. MacEachren, and Biliang Zhou. 2005. Multivariate Analysis and Geovisualization with an Integrated Geographic Knowledge Discovery Approach. *Cartography and geographic information science* 32: 113-32.

Guthrie, John T., Shelley Weber, and Nancy Kimmerly. 1993. Searching documents: cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary educational psychology* 18: 186-221.

Harris, Robert L. 1999. *Information graphics: a comprehensive illustrated reference.* Atlanta, GA: Management graphics.

Hayes, Conor, Paolo Avesani, and Uldis Bojars. 2006. An analysis of bloggers, topics and tags for a blog recommender system. Edited by Bettina Berendt, Andreas Hotho, Nunia Mladenic and Giovanni Semeraro. *From Web to Social Web: Discovering and Deploying User and Content Profiles: Workshop on Web Mining, WebMine 2006.* Berlin, Germany, pp. 1-20.

Heath, Lenwood S., Deborah Hix, Lucy T. Nowell, William C. Wake, C.A. Averboch, Eric Labow, Scott A. Guyer, Dennis J. Brueni, Robert K. France, Kaushal Dalal, and Edward A. Fox. Envision: a user-centered database of computer science literature. *Communications of the ACM* 38n4: 52-53.

Hill, Linda L. 2006. *Georeferencing : the geographic associations of information.* Cambridge, MA: MIT Press.

Hill, Linda L., Olha Buchel, Greg Janee, and Marcia L. Zeng. 2002. Integration of Knowledge Organization Systems into Digital Library Architectures: Position Paper. *Thirteenth ASIS&T SIG/CR Workshop on Reconceptutalizing Classification.* Philadelphia, PA.

Hubmann-Haidvogel, Alexander, Arno Scharl, and Albert Weichselbraun. 2009. Multiple Coordinated Views for Searching and Navigating Web Content Repositories. *Information sciences* 179n12: 1813-21.

Imhof, Eduard. 1975. Positioning names on maps. *The American cartographer* 2n2: 128-44.

Jones, Michael T. 2008. 2008 Google I/O Session Videos and Slides. *The world's information in context*. http://sites.google.com/site/io/the-worlds-information-in-context (accessed July 15, 2008).

Kang, Hyunmo, and Ben Shneiderman. 2000. Visualization methods for personal photo collections: browsing and searching in the PhotoFinder. *2000 IEEE International Conference on Multimedia and Expo*, pp. 1539-42.

Kapler, Thomas, and William Wright. 2004. GeoTime Information Visualization. *IEEE Symposium on Information Visualization (InfoVis 2004)*, pp. 25-32.

Keim, Daniel A., Stephen C. North, Christian Panse, and Jörn Schneidewind. 2003. Visualizing geographic information: VisualPoints vs CartoDraw. *Information visualization* 2n2: 58-67.

Kohonen, Teuvo. 1995. *Self-organizing maps.* Berlin; Heidelberg; New-York: Springer Verlag.

Koua, Etien Luc. 2005. Computational and Visual Support for Exploratory Geovisualization and Knowledge Construction. Thesis, Faculty of Geographical Sciences, Utrecht University, Utrecht, The Netherlands.

Lamping, John, Ramana Rao, and Peter Pirolli. 1995. A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. Edited by I. Katz, R. Mack, L. Marks, M.B. Rosson and J. Nielsen. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems.* Denver: ACM, pp. 401-08.

Le Grand, Bénédicte, and Michel Soto. 2002. Visualisation of the Semantic Web: Topic Maps Visualisation. *IEEE IV 2002.* London.

Leazer, Gregory H, and Jonathan Furner. Topological indices of textual identity networks. *Knowledge: Creation, Organization and Use: Proceedings of the 62nd American Society for Information Science Annual Meeting*, pp. 345-58.

Li, Xia, and Menno-Jan Kraak. 2008. The Time Wave A new method to visually explore geo-data in time space. *Cartographic journal* 45n3: 193-200.

Liao, Ke, Diansheng Guo, Alan M. MacEachren, and Jin Chen. 2006. A Visualization System for Space-Time and Multivariate Patterns (VIS-STAMP). *IEEE transactions on visualization and computer graphics* 12: 1461-74.

Library of Congress. n.d. *Library of Congress.* http://www.loc.gov/catdir/cpso/lcco/ (accessed October 25, 2007).

Library of Congress Network Development and MARC Standards Office. 2007. *Introduction.* http://www.loc.gov/marc/languages/ (accessed July 15, 2008).

Lohmann, Steffen, Jurgen Ziegler, and Lena Tetzlaff. 2009. Comparison of tag cloud layouts: Task-related performance and visual exploration. Edited by Tom Tom Gross, et al. *Human-Computer Interaction - INTERACT 2009: 12th IFIP TC 13 International Conference.* Uppsala, Sweden, pp. 392-404.

Lowe, Richard K. 2003. Animation and learning: Selective processing of information in dynamic graphics. *Learning and instruction* 13: 157-76.

Lowe, Richard K. 1999. Extracting information from an animation during complex visual learning. *European journal of psychology of education* 14: 225–44.

MacEachren, Alan M. 1995. *How maps work: representation, visualization, and design.* New York, NY: The Guildford Press.

MacEachren, Alan, Xiping Dai, Frank Hardisty, Diansheng Guo, and Gene Lengerich. 2003. Exploring High-D Spaces with Multiform Matricies and Small Multiples. *Proceedings of the International Symposium on Information Visualization*.

Moen, William E., and Shawn D. Miksa. 2007. Informing the Future of MARC: An Empirical Approach. *MARC Content Designation Utilization: Inquiry and Analysis.* http://www.mcdu.unt.edu/wp-content/ALA2007ProgramHandout22June2007.pdf (accessed October 10, 2007).

Monroy, C., R. Kochumman, R. Furuta, and E. Urbina. 2002. Interactive timeline viewer (ItLv): A tool to visualize variants among documents. In *Visual interfaces to digital libraries*, edited by K., Chen, C. Börner. Springer-Verlag, pp. 39-49.

Montello, Daniel R., Sarah I. Fabrikant, Marco Ruocco, and Richard Middleton. 2003. Spatialization: testing the First law of cognitive geography on point-spatialization displays. *Spatial Information Theory: Foundations of Geographic Information Science. Conference on Spatial Information Theory (COSIT) 2003, Lecture Notes in Computer Science.* Berlin, Germany: Springer Verlag.

Moriyama, Masaki, Delwyn L. Harnisch, and Shinichi Matsubara. 1994. The development of graphic symbols for medical symptoms to facilitate communication between health care providers and receivers. *Tohoku journal of experimental medicine* 174: 387-98.

Newton, Douglas P. 1994. Supporting the Comprehension of Tabulated Data. *Journal British educational research journal* 20: 455-63.

Nowell, Lucille Terry. 1997. Graphical Encoding for Information Visualization: Using Icon Color, Shape, and Size To Convey Nominal and Quantitative Data. Thesis, Faculty of the Virginia Polytechnic Institute and State University.

Peuquet, Donna J. 2002. *Representations of space and time.* New York, NY: The Guildford Press, 2002.

Pickett, Ronald M., and Georges G. Grinstein. 1988. Iconographic Displays for Visualizing Multidimensional Data. *Proceedings of the 1988 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 514-19.

Plaisant, Catherine. 2005. Information visualization and the challenge of universal usability. In *Exploring geovisualization*, edited by Alan M. MacEachren, Menno-Jan Kraak and Jason Dykes. New York, NY: Elsevier Ltd, pp. 53-82.

Rensink, Ronald A., J. Kevin O'Regan, and James J. Clark. 1997. To see or not to see: The need for attention to perceive changes in scenes. *Psychological science* 8: 368-73.

Reuter, Kara. 2007. Assessing aesthetic relevance: Children's book selection in a digital library. *Journal of the American Society for Information Science and Technology* 58: 1745-63.

Rosario, Geraldine E., Elke A. Rundensteiner, David C. Brown, Mathew O. Ward, and Shiping Huang. 2003. Abstract Mapping Nominal Values to Numbers for Effective Visualization. *InfoVis2003*.

Savolainen, Reijo, and Jarkko Kari. 2006. User-defined relevance criteria in web searching. *Journal of documentation* 62: 685.

Sedig, Kamran, Rowhani, Sonja, and Liang, Hai-Ning. 2005. Designing interfaces that support formation of cognitive maps of transitional processes: An empirical study. *Interacting with computers: the interdisciplinary jr. of human-computer interaction* 17: 419-52.

Sedig, Kamran, and Mark Sumner. 2006. Characterizing interaction with visual mathematical representations. *International journal of computers for mathematical learning* 11n1: 1-55.

Sheehy, Kieron. 2002. The effective use of symbols in teaching word recognition to children with severe learning difficulties: A comparison of word alone, integrated picture cueing and the handle technique. *Journal of disability development in education* 49n1: 47-59.

Shneiderman, Ben. 1992. Tree visualization with treemaps: A 2-D space-filling approach. *ACM transactions on graphics* 1n1: 92-99.

Sieber, René, Christoph Schmid, and Samuel Wiesmann. 2005. Smart Legend - Smart Atlas!" *Proc. of the 22th Int. Conference of the ICA.* A Coruña, Spain.

Slocum, Terry A., Robert B. McMaster, Fritz C. Kessler, and Hugh H. Howard. 2005. *Thematic cartography and geographic visualization.* Upper Saddle River, NJ: Prentice Hall.

Smiraglia, Richard P., ed. *Works as entities for information retrieval.* Binghampton, N.Y.: Haworth Information Press, 2002.

Spence, Robert. 2008. *Information visualization: design for interaction.* 2 nd. Essex, England: Pearson Education Limited, 2007.

Stryker, Michael, Ian Turton, and Alan MacEachren. Health GeoJunction: Geovisualization of news and scientific publications to support situation awareness. *Geospatial Visual Analytics 2008.* Park City, Utah.

Svenonius, Elaine. 2000. *The intellectual foundation of information organization.* Cambridge, Mass.: MIT Press.

Tang, Diane, Chris Stolte, and Robert Bosch. 2003. Design Choices when Architecting Visualizations. *Proceedings of the Ninth IEEE Symposium on Information Visualization*.

Tomaszewski, Brian. 2008. Producing Geo-historical Context from Implicit Sources: A Geovisual Analytics Approach. *The cartographic journal* 45: 165-81.

Tombros, Anastasios, Ian Ruthven, and Joemon M. Jose. 2005. How users assess web pages for information seeking. *Journal of the American Society for Information Science and Technology* 56: 327-44.

Tremmel, Lothar. 1995. The Visual Separability of Plotting Symbols in Scatterplots. *Journal of computational and graphical statistics* 4: 101-12.

Tufte, Edward. 1990. *Envisioning information.* Cheshire, CT: Graphics Press

Tufte, Edward. 2001. *The visual display of quantitative information.* 2nd. Cheshire, Connecticut: Graphics Press.

Vasiliev, Irina R. 1997. Mapping time: discussion of historical attempts to define time, and integration of time information into diagrams of geographic data. *Cartographica* 34n2: 1-44.

Veerasamy, Aravindan. 1997. Visualization and User Interface Techniques for Interactive Information Retrieval Systems. Unpublished PhD thesis, Georgia Institute of Technology, Atlanta.

Vonk, Wietske, and Leo G.M. Noordman. 1990. On the control of inferences in text understanding. In *Comprehension processes in reading*, edited by David A. Balota, Giovanni B. Flores d'Arcais and Keith Rayner, 447-464. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ware, Colin. 2004. *Information visualization: perception for design.* San Francisco, CA: Elsevier Inc.

Weerdt, David De, Rani Pinchuk, and Richard Aked. 2007. TopicMaker – An Implementation of a Novel Topic Maps Visualization. Edited by L. Maicher, A. Siegel and L. M. Garshol. *Leveraging the Semantics of Topic Maps – Second International Conference on Topic Map Research and Applications, TMRA 2006.* Berlin, Heidelberg, New York: Springer Verlag, pp. 33-43.

Wilhelm, Adalbert. 2008. Linked Views for Visual Exploration. In *Handbook of Data Visualization*, edited by Chun-houh Chen, Wolfgang Härdle and Antony Unwin, 199-215. Berlin: Springer.

Winn, William. 1987. *Charts, graphs, and diagrams in educational materials.* Vol. 1, in *The Psychology of illustrations: instructional issues*, edited by Harvey A. Houghton and Dale M. Willows, 152–198. New York: Springer-Verlag.

Wise, James A., James J. Thomas, Kelly Pennock, David Lantrip, Marc Pottier, Anne Schur, and Vern Crow. 1995. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. *Proceedings IEEE Visualization 95.* Los Alamitos, CA: IEEE Computer Society Press, pp. 51-58.

Xu, Yunjie (Calvin), and Zhiwei Chen. 2006. Relevance judgment: What do information users consider beyond topicality?" *Journal of the American Society for Information Science and Technology* 57: 961-73.

Zeng, Marcia L. 2008. Knowledge Organization Systems (KOS). *Knowledge Organization* 35: 160-82.

Zerneke, Jeannette L., Michael K. Buckland, and Karl Carl. 2006. Temporally Dynamic Maps: The Electronic Cultural Atlas Initiative Experience. *Human IT* 8n3: 83-94.