

Support for the Learner: What, Where, When, and Who

Final Report

on the Institute of Museum and Library Services
National Leadership Grant for Libraries, no. LG-02-04-0041-04,
Oct 1, 2004 – Dec 31, 2006.

Michael Buckland, Fredric C. Gey and Ray R. Larson,
Electronic Cultural Atlas Initiative,
University of California, Berkeley, CA
March 30, 2007

INTRODUCTION

Learning, if it is to be more than memorizing, requires an understanding of context. Anyone setting out to learn a new topic would, traditionally, take a pad of paper, start with resources available at home, visit the reference area of the local library, then move on to a steadily widening circle of resources: the library stacks, other libraries, museums, archives, local, state or federal government agencies, cultural and historical societies, newspapers, experts, and so on, depending on the topic and the purpose. The library reference collection is especially convenient for addressing the What?, When?, Where?, Who?, Why?, and How? because it has carefully selected specialized resources. Dictionaries, encyclopedias, and the library subject headings help clarify the topic and its terminology, and the events, dates, institutions, and persons involved. Biographical dictionaries describe individuals and their life activities. Chronologies and newspaper indexes pinpoint pertinent dates. Gazetteers and atlases show where places are. Associating topics with persons, times, and places is especially important when learning about the arts, history, and cultural heritages. Bibliographies, catalogs, and directories lead to further resources in other collections and in different archives, libraries and museums.

A networked environment greatly increases the range and variety of accessible resources, but digital library and museum environments have yet to provide learners with a structured environment analogous to the traditional library reference collection. The digital environment is still weak in providing that kind of “intermediate infrastructure” in the zone between the learner and the best resources. More assistance is needed for navigating multiple metadata, building crosswalks between different vocabularies, and integrating search results into personal computing environments.

This project set out to demonstrate how intermediate infrastructure could be developed to provide access to the best available resources with the kind of supportive learning environment that one associates with a reference library. Even modest improvements in standards and best

practices will, individually and cumulatively, improve our collective ability to find materials of different kinds related to any given topic. This report summarizes the work done. Fuller details can be found on the project website: <http://ecai.org/imls2004>, especially in the list there of related publications.

The project was inspired by two design challenges. One was educational: to consider K-12 history and social science teaching and to ask what kind of search support would enable a teacher to find additional resources explaining the background on any topic, person, institution, or event mentioned in an assigned textbook. The other was museum-related: For any given museum object, what kind of search support would facilitate discovering contextualizing resources elsewhere in archives, databases, libraries, or museums? What else is known to come from the same place and/or time, or having the same purpose? What literature, archival records, images, sound recordings, or other documents relate to it specifically or, more generally, to the kind of object it is? With both design challenges the task is to determine what courses of action might make the digital apparatus of metadata, search engines, and interoperability more supportive for learners. Figure 1 shows a project diagram.

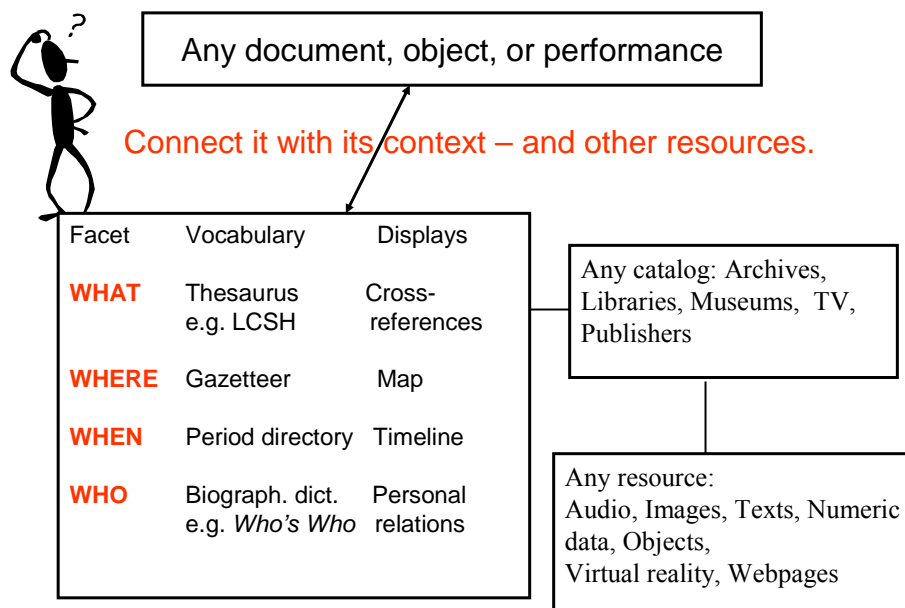


Fig. 1: Project diagram.

It became apparent early in the project that these two design challenges were too ambitious for the time and resources available, so, with the concurrence of IMLS, we worked on more elementary steps that would help to build the bases for eventually addressing these and similar challenges in a practical way.

WHAT – Mapping between vocabularies

Most well-edited resources have some kind of indexing or categorization and it is widely understood and accepted that the inherent ambiguity and instability of language needs to be

“controlled” in indexes, so, for example, synonyms should be explicitly related and hierarchical and other relationships indicated. However, the “vocabularies” of categories, codes, and indexing terms used vary widely and can be quite complex. In practice, efficient, effective search and selection requires some familiarity with whatever vocabulary is being used in the resource being searched.

Getting from the words that the learners use (the query vocabulary) to the terms in the metadata (entry vocabulary) is fundamental. How is anybody looking for ALIENS FROM OUTER SPACE to know that they should have been searching under EXTRATERRESTRIAL BEINGS and LIFE ON OTHER PLANETS? Who could guess the US Patent Classification for PEANUT BUTTER?

“Relative indexes” (to use Dewey’s term) leading from the terms and phrases of users to the most apt choices in the topical metadata of the resource to be searched are indexes from the *Query vocabulary* to the *Entry vocabulary*. Learners also need help in moving between entry vocabularies. If you did know that the US Patent Classification number for peanut butter is 426/633.00, how easily would you find the corresponding International Patent Classification code: A23L 1/38? The Unified Medical Language System (UMLS) of the National Library of Medicine is the outstanding example of mapping across vocabularies, but very expensive to extend or replicate. These kinds of mappings are especially needed in a digital environment where schedules are not so easily arrayed for browsing.

In a Web environment multiple resources use quite different descriptive metadata vocabularies, so mapping between related terms in *different* vocabularies becomes very important, but has received relatively little attention. Manually relating the terms in two or more vocabularies becomes extremely labor-intensive as well as being inherently obsolescent as each vocabulary evolves. There are, however, techniques using statistical association and natural language processing which can generate inexpensive mappings if a corpus of records is available as a training set (Buckland, Chen, Chen & others, 1999; Buckland, Gey & Larson, 2002; Buckland, Chen, Gey & others, 2006).

An important and neglected aspect of search support arises when different learners ask for literature on the same topic but, because they have different interests and backgrounds, they do not, in fact, have quite the same need. A geriatrician, an anesthesiologist, and a surgeon, for example, might all ask for literature on, say, cardiac arrest, but because their interests in cardiac arrest differ they would not, in fact, want to be given the same literature. Effective search support needs to be designed to lead to different resources accordingly, e.g. Find literature on “Cardiac arrest” *suitable for a geriatrician*. Vivien Petras, graduate student researcher on this project, addressed this issue in her doctoral dissertation *Translating dialects in search: Mapping between specialized languages of discourse and documentary languages*, which received the outstanding doctoral dissertation of the year award of the American Society for Information Science and Technology (Petras 2006).

Another innovation was our creation of a mapping between WHAT in a gazetteer and WHAT in library subject catalog. We took the Geographic Description Codes (aka “Feature types”) used by the National Geo-intelligence Agency and mapped them to the corresponding Library of Congress Subject Headings. (More below).

Naming WHAT – marking documents with descriptive names and assigning documents to named categories – is culturally based and necessarily unstable over time, undermining the effectiveness and stability of retrieval systems. It is the site of tensions between the procedural

need for stable marks and the inherent multiplicity and instability of linguistic expressions used to represent topics. Our theorizing of the nature of the issues, tensions, compromises, and inherent obsolescence involved led to an invited chapter in a linguistics book (Buckland 2007b).

WHERE – Place, space, and gazetteers

Place names are notoriously ambiguous (e.g. which of the cities named Lafayette?), multiple (St. Petersburg, Санкт-Петербургский, Saint-Petersbourg, etc.), unstable (e.g. St. Petersburg 1703-1914; Petrograd 1914-1924; Leningrad 1924-1991; St. Petersburg, again, in 1991), and/or vaguely defined (e.g. The Midwest). Geopolitical entities are additionally unstable because boundaries and political structures change: Think of local government boundaries, the Balkans, or the widely varying boundaries, over the years, of Poland. Improved “best practices” could substantially improve support for geographical searching. It is useful to distinguish *place*, a cultural concept, from *space*, a physical construct. Place name gazetteers are best known from their appearance as large pages of small type at the back of atlases, where they also serve as indexes to the maps. But gazetteers exist in their own right as, in effect, bilingual dictionaries linking places and spaces. Gazetteers, in library terminology, are place name authority files. Latitude and longitude can disambiguate the numerous Lafayettes and show the identity of different names, e.g. Beijing and Peking. The importance of gazetteers is that they allow named places to be found or represented on a map.

Museums and others who deal with historical material have the additional problem that the common practice of using the names and boundaries of today’s political jurisdictions may make little sense when dealing with past times. Support for geographic search could be improved if a few additional steps were taken:

1. Authority lists of place names, as found in libraries, should either have geographical coordinates (latitude and longitude) added, effectively making them into gazetteers, or, better, be linked to authoritative gazetteers maintained elsewhere.
2. Gazetteer entries should (but rarely do) include an indication of *when* that name was in use;
3. Now that catalogs are no longer made and presented in card form but by computer interface, there is no reason not to provide map interfaces. The geographical distribution of retrieved records can be shown as an aid to a more refined selection. Also, drawing a region of interest on a map interface is a convenient way of expressing the geographical scope of a search: Find all ceramic items in the collection from *this area*.
4. Understanding and explaining change over time is important. Maps on paper have limited ability to show chronological changes. But digital maps can show changes over time dynamically (Zerneke, Buckland & Carl, 2006).
5. Gazetteer entries use geographical description codes (aka Feature types) to indicate the kind of place named: Castle, lighthouse, lake, city, etc. These feature type codes can be linked with corresponding subject headings.

A comparison of the National Geo-intelligence Agency’s Geographical Description Codes (GDC) with Library of Congress Subject Headings (LCSH) revealed differences in style, emphasis, scope, and scale. Nevertheless, in most cases there are sensible matches. For example, the NGA code “School” means a school building, and corresponds to LCSH “School buildings” (for the physical feature) and to LCSH “Schools” (for schools as institutions). Mapping between

these two vocabularies allows one to move from the literature on, say, the *topic of lighthouses* to locating (through a gazetteer and a map) *instances of lighthouses* on the ground. Moving in the other direction, if you find an actual lighthouse, you could search for literature about that particular lighthouse, lighthouses in that region, or lighthouses generally. This extended the work done in our prior IMLS-supported project “Going Places in the Catalog: Improved Geographical Search” (ecai.org/imls2002/).

WHEN – Use of events to denote time

Clocks and calendars provide the obvious way to measure and record time, but in both speech and in writing, we find that events, rather than calendar dates, are widely used to denote points or periods of time. Events are commonly wars (“Civil war,” “World War II”), reigns, dynasties, and administrations (“a Louis XIV clock,” “under Clinton”), cataclysmic events (“after the Lisbon earthquake”) or personal (“after graduation”). This use of events to denote time tends to be situational, multiple, ambiguous, and unstable. A “civil war weapon” would date from the seventeenth century in England or the twentieth century in Spain. “The Great War” was renamed the “First World War” and, in some quarters, the Vietnam War is coyly referred to as the “Vietnamese Conflict.”

Directories of named time periods that link between events used in common language as temporal indicators with precise dates (as used in most information systems) are not yet a developed genre. So, since the characteristics of named time periods resemble those of place names, we developed an analogous solution. We drew on the chronological subdivisions of Library of Congress Subject Headings (\$y) as examples and created a named time period directory modeled on the design of place name gazetteers (Petras, Meiske, Larson, Zerneck, Carl & Buckland, 2005; Petras, Larson & Buckland, 2006).

For each named period, a code for type of period (war, dynasty, cataclysm, etc.) and corresponding calendar dates were provided, and, just as a gazetteer should specify when a name was in use, so also a time period should indicate the geographical context, thus:

Place name gazetteer

Place name–Type--Geographical markers (lat. & long)--When in use

Named period directory

Name of period–Type–Chronological markers (calendar)–Where used

Just as gazetteers support maps, time period directories support time lines. Just as a map allows one to see what else is at or near the same place, time periods and chronologies enable one to see what else was happening at, or about, the same time.

WHERE and WHEN: Geo-temporal Metadata

Gazetteers and Time period directories are related. Just as gazetteer entries have a period aspect, time periods often have a geographical aspect. The inclusion of *both* place and time aspects in *both* gazetteers and time period directories allows geo-temporal relationships: What else was near this place around that time? Descriptive metadata systems can be seen as a form of infrastructure and these kinds of links facilitate the construction of metadata infrastructures (Buckland, 2006).

WHO – Names and activities

We had approached WHO with modest intentions, seeing the task as primarily a matter of forwarding searches to find entries in biographical dictionaries, to biographical records in encyclopedias, and to library catalogs to find books by and books about individuals. Fortunately, the desirability of disambiguating different persons with the same name and of connecting different names for the same person is well-understood and widely implemented.

However, we realized that there were some major difficulties and very rich possibilities. One very simple matter is that there is no agreed standard for specifying even a simple name: *Smith, John* in the library is *John Smith* outside the library and unlikely to be found using “Smith, John” as a query. We encountered a lack of established vocabulary practices for expressing interpersonal relationships (e.g. mentor, classmate, colleague, opponent) outside the basic family relationships such as parent, child, sibling, and spouse. Also, the range of network-accessible resources that one can reach using federated search techniques (such as Z39.50) is still small.

More importantly, biographical texts are very richly descriptive, e.g.

Emanuel Goldberg. Born Moscow 1881. PhD under Wilhelm Ostwald, University of Leipzig 1906. Director, Zeiss Ikon, Dresden, 1926-1933. Moved to Palestine 1937. Died Tel Aviv 1970.

Almost every element could be expanded: Moscow, Dresden, Palestine, and Tel Aviv in a gazetteer; Ostwald in a biographical dictionary; Zeiss Ikon through a business directory; “Born,” “PhD,” “Director,” “moved,” and “died” indicate actions or achievements. The dates can be used to learn what else was happening at that time in that place or topic. For example, a library catalog search on Leipzig (aka Leipsic!) qualified by publication date 1900-1910 can lead to numerous contemporary descriptions of what Leipzig and its University were like then.

Although biographical text is important in many contexts, the standards and best practices for representing what people do and the events in people’s lives are seriously inadequate. (For a useful survey revealing how inadequate see Text Encoding Initiative Consortium. (2006). *Report on XML mark-up of biographical and prosopographical data*. <http://www.tei-c.org/Activities/PERS/persw02.xml>).

We also formulated a promising approach to representing an individual’s activities by expressing life activities as a set of separate activities (or events) at whatever level of detail is desired, and then encoding each activity as a 4W-tuple of *what* kind of activity, *where* it took place, *when* and for how long it occurred, and *who* else was involved. This approach was, in effect, a microcosm of the whole project. Fortunately, we are able to address the challenges to making better use of biographical texts in a separate subsequent project, also supported by an IMLS National Leadership Grant, *Bringing Lives to Light: Biography in Context* (<http://ecai.org/imls2006>). The hope is that this approach could be generally acceptable across different communities and that the vocabularies already established within each community could be used to express *what* (topical subject headings), *where* (place name gazetteers), *when* (named time period directories), and *who* (name authority files and biographical dictionaries). If successful the basis for a great deal of interoperability within and between communities could emerge.

SEARCH SUPPORT

So far we have been concerned with improvements in description. Searching can also be facilitated.

Text mark-up commonly incorporates embedded links to related sources which provide further explanation or validation. At best these links are obsolescent as newer publications appear. An alternative is to provide a dynamic link in the form of a search query. As an example, *ECAI Iraq* (<http://ecai.org/iraq/>), a temporal-spatial portal into existing digital resources about the history, cultural sites, archaeological excavations, and heritage preservation initiatives relating to Iraq, contains a series of web pages for individual historic sites (Electronic Cultural Atlas Initiative, 2003). Clicking on the first three links on each site page automatically generates search queries, using the Z39:50 search and retrieve protocol, for material concerning that site in the library catalogs of the University of California, the research libraries of the United Kingdom, and the Library of Congress. The merit of this approach is that the material retrieved will be as up-to-date as those libraries' cataloging rather than entries from a static, obsolescing bibliography.

The structures of URLs are usually opaque and of little meaning, but clearly structured URLs can greatly simplify searching from remote locations. The Timeline of Art History section of the Metropolitan Museum website (<http://www.metmuseum.org/toah/>) is an excellent example of helpful design. This section uses eleven defined time periods (from "20,000-8,000 BC" through "1900 AD – present") and, for places, a hierarchical structure starting with nine major regions. Each time period and each geographical area has a simple, easily-discerned code, clearly visible in the /toah URLs. Knowing these codes, they can be mapped to the categories for time and place in any local system and inserted algorithmically into a link. If one were interested in the art of southern India during in the eighth century CE and knew that it would be categorized at the Met website by "06" for the time period 500-1,000 AD and "sss" for South South Asia, one can insert "sss" and "06" into the /toah URL extensions to form <http://www.metmuseum.org/toah/ht/06/sss/ht06sss.htm> which, *for this topic*, is a good starting point for searching for the Met's rich collections. Mapping local metadata to the /toah codes and inserting them in a link allows one to go *directly* from inside any local Web page to the appropriate page *within* the Met's website.

The *Wikipedia* provides a similarly useful design for biographical searching. Creating a URL by adding "Firstname_surname" to the stem <http://en.wikipedia.org/wiki/> to form, for example, http://en.wikipedia.org/wiki/John_smith will lead directly to any biographical article(s) for person(s) with that name in the *Wikipedia*. As can be seen from these examples, strategies for linking different resources to each other depend in the individual structures of the resources in question. Consequently, it should always a goal to standardize metadata structures to make these linking processes easier to generate and to automate.

INTERFACES AND PROTOTYPES

Designing a user-friendly interface capable of supporting searches along four dimensions proved a challenge. Not only do four different kinds of search need to be supported, but there also needs to be provision searches *combining* two or more, for example, time and place – and support access to diverse media forms – all in one intuitive, easy-to-use interface. After some delay we decided to develop two prototypes with different characteristics.

The “4W Interface”

The primary “4W interface” is a web system with dynamic content generation. The content displayed on a page is dependent on context and is generated on the fly when a link is chosen. However, the system is stateless, so information about the user and their previous choices is not collected or used in navigation management. The 4W Interface includes three major components: natural language search, menu driven browse, and dynamic map browse and query.

The searches in the 4W Interface take advantage of “entry vocabulary” systems developed in previous projects and federated search links. The search starts when a user enters a text query and the interface displays a list of related topics and a variety of searchable target resources are offered. The results may also include embedded references to full text resources and to related Wikipedia searches.

The browse function enables discovery by context dimension. The interface provides menus, which display information if data is available. This section of the interface provides connections into a wide variety of reference resources enabled by crosswalk reference tables. The reference tables include the Time Period Directory which links event names with Western calendar dates (ISO standard) and a geographic reference. The geographic references were simplified to country, state, or major city. This method is used because a standard global historical gazetteer does not yet exist.

Connections to the reference works used are easy to maintain because the reference works themselves maintain a structured resource. They are either using explicit codes to reference locations or have a structured website where URLs are consistently built using discoverable codes. A primary reference or crosswalk table links locations with codes used for them by a number of major information providers.

The browse is initiated by either location or time, then further dimensions are offered such as events, or people. This single dimensional browse was used since the results are a matrix where the more dimensions you choose the less likely you are to find results, so, a guided approach is used.

The results of the browse are resources that are external to the 4W Interface, including dynamic searches of the Library of Congress catalog using the Cheshire search service and queries into the Wikipedia, Metropolitan Museum of Art, and others. Further refinement of the query and further browsing can take place in the external website. However, this information is not captured by the 4W system. Further discussions about approaches to capturing refinements and found resources are ongoing.

Map display and query use the TimeMap Java browser and are based on previous ECAI models. This section of the 4W Interface supports display and query by two dimensions, time and place, simultaneously. Using the interactive TimeMap it is possible for a user to discover what information is available by location and the user can select a time range using the time slider bar. Then a two dimensional query is sent to the reference database and webpage results are built for the resulting time and place. Once the user has found this result set, the option is given to the user to broaden the search either by time or location, returning to the single dimensional browse function.

The system architecture for the interface includes: a website with embedded java scripts hosted by ECAI and by International and Area Studies; relation databases for the reference tables such as the time period directory developed for the project and the cross walk tables;

programming to generate dynamic web page content in perl and asp; Cheshire search scripts; dynamic maps generated using ECAI's TimeMap dynamic map development tools and displayed using the online Java version of TimeMap.

The interface which is openly available at <http://ecai.org/imls4w/> Recommended beforehand is a powerpoint presentation that introduces the project and then the interface available at <http://ecai.org/imls2004/bucklandrot.ppt>. (This invited presentation to the Digital Heritage Netherlands annual conference in Rotterdam is based on our presentation to the IMLS Board on September 28, 2006.) Based on our experience with the "4W Interface" we are currently developing an alternative design for the "Bringing Lives to Light" project.

Figure 2 illustrates one navigational point in the "4W Interface".

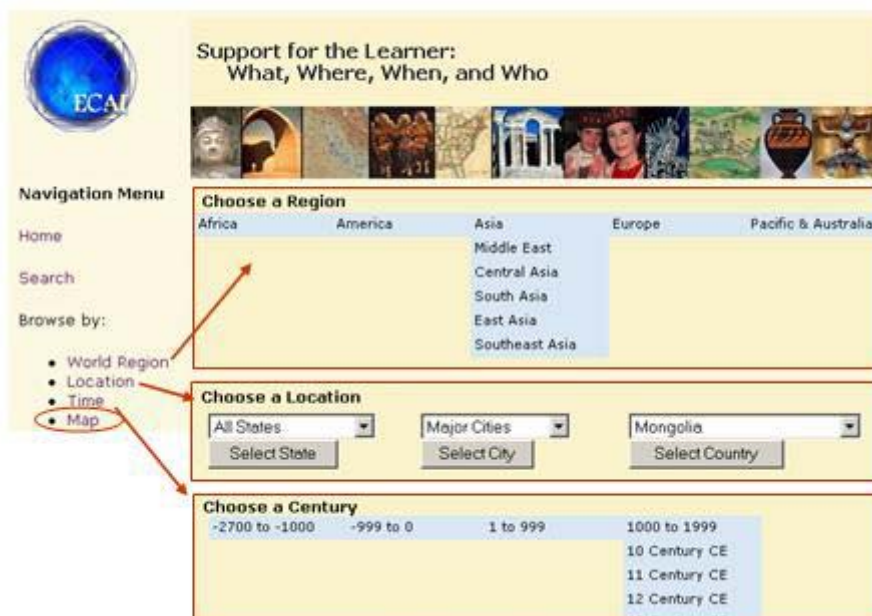


Fig 2. The Navigation Menu page for the 4W Interface. This composite image shows four search options indicated by red lines: Browse by region; search by location using drop-down menus; drill down into the time period directory; or click to an interactive map interface.

The "4W Facts Interface"

Since our earlier IMLS-supported project "Seamless Search of Text and Numbers" we have continued to experiment with interfaces to support search of both text corpora and socio-economic numeric data sets. We wanted to incorporate this functionality into the "4W Interface," but since available numeric datasets did not match the wide-ranging geo-temporal range of the latter, we decided we would learn more by developing a separate "4W Facts Interface" focused on California U.S. local history and using, as resources, *American Factfinder* (Census Bureau, <http://factfinder.census.gov>), *Counting California* (a large compilation of statistical tables made available by the California Digital Library, <http://countingcalifornia.cdlib.org>), and the University of California online library catalog. With these resources a unified search support is

provided for geographical, temporal and/or topical searches for current or historical information – factual, statistical, and bibliographical – for localities, including census data back to 1790. The scope has been extended to Delaware, New Jersey, and Pennsylvania. See Fig. 3. A detailed technical report is in preparation.

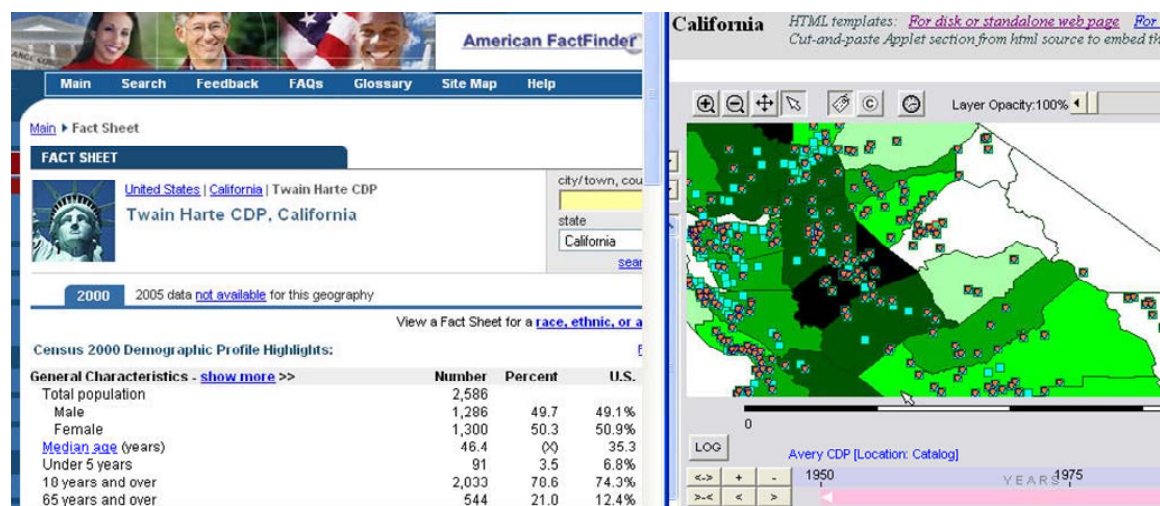


Fig. 3: The “4W Facts Interface”: Screen shots of geo-temporal search interface (right) and census data (left) for the small town of Twain Harte, CA.

DOSSIERS

Harmonizing the excerpts from heterogeneous resources into a coherent dossier is a challenge. METS (the Metadata Encoding and Transmission Standard) offers a solution and we envisaged using METS to enable learners to create dossiers in the form of an ad hoc web portal. This aspect we gave a low priority and did little work on it, regarding it as subordinate to needs assessment and interface design, both of which proved more complex and time consuming than expected.

NEEDS ASSESSMENT AND EVALUATION

At a narrowly technical level, proof of concept was provided by the prototype search portal at <http://ecai.org/imls4w/>

This was an exploratory project and potential users' needs were assessed in an exploratory way. Instead of the usual practice of introducing a service and then surveying its users, we sought to find out about needs without direct reference to any intended prototype. At Berkeley an appeal for faculty volunteers was sent out to all instructors teaching courses relating to American cultures and interviewed them through in-depth, largely undirected discussions. At Dominican University of California volunteers were invited from faculty, especially those concerned with courses in Education. Also, at Dominican, campus librarians were involved actively but informally with the project.

Although the formal project Outcomes were achieved, the deeper issues revealed were sobering. Initial interviews presented alarming and unexpected challenges to the overall project assumptions. We found a deeper chasm than expected between the realities of how faculty cope

and the expectations of providers of digital library resources. Also, a significant difference emerged between the usual evaluation practices of librarians (How is the service being used?) and the more individualized evaluation practices of educational researchers (What helped this person learn?). Faculty work on a semester cycle and IMLS approved a no-cost three month project extension through the final Fall semester, to December 31, 2007, to allow us time for additional discussion and analysis.

Participants had a hard time articulating precise information needs. In general, faculty tended to make do with what they knew of, even when they suspected that there were valuable information sources going overlooked and unexplored. Few were able to make connections between what they were interested in topically and the type of information they hoped to find. Faculty tend to “satisfice.” There was little difference between Berkeley and Dominican.

We found that faculty and end-users are reluctant to engage in collaborative critical reflection on their information use and were primarily concerned with immediate needs -- and that information professionals tend not to work within a framework of instruction and have perspectives different from their users.

The demonstrations of the prototype interface were met with enthusiasm by all faculty and student participants but many had a difficult time understanding the purpose of the interface. Alarmingly, many did not understand key concepts in information organization and retrieval, such as classification schemes. While many were fascinated, they could not see how the interface would help them specifically accomplish their work. Searchers expect immediate results and were unwilling to browse. Facilitated discussion about how to use the interface during the demonstration illustrated the need for concrete guiding and coaching on interface use. This point was also upheld by interviews conducted by students who found that users tended not to read helpful tips printed on web pages but did respond to verbal feedback

Overall, we were impressed by the low level of engagement with online resources, modest information technology skills, and weak motivation to change. A fuller report and discussion of the needs assessment and evaluation is in preparation.

DOCUMENTATION AND DISSEMINATION

The project website -- <http://ecai.org/imls2004> – makes project documentation available, including the “4W Interface” prototype and a list (with links) to more than twenty related papers and presentations, some at major conferences, others in leading professional and technical journals.

Project work was described and disseminated in numerous presentations at conferences, seminars, and invited talks in at least eight states of the USA and also abroad (China, France, Netherlands, and the U.K.).

The Advisory Committee was valuable for communication both to and from the project.

THE BIGGER PICTURE

Our work opened up and to some extent addressed some major issues of wider import which went beyond the scope of this project:

1. *The teaching library*: The chasm between the realities of faculty life and the service assumptions of library service providers and the disconnect between educators and librarians with respect to “information literacy” were larger than we had expected and larger than the

professional literature reveals. The implications are profound. Much deeper dialog, mutual understanding, and collaboration between educators and librarians are called for, and the barriers to the substantial use of additional digital library services imposed by faculty satisfying behavior are high in deed.

2. *Time, time periods, and search.* In prior work we have shown how geographic search can be greatly improved through careful attention to the relationships between place names, spatial georeferencing (latitude and longitude), place name gazetteers, and map displays. We now believe that a comparable advance in temporal search through careful attention to named time periods, named events used to mark time, calendar time, time period directories, time lines and chronologies is emerging.

3. *Differentiated support for learners in different domains.* Before libraries adopted digital techniques there was little scope for supporting searches in the form “Find literature on ‘Cardiac arrest’ suitable for a geriatrician. It has been shown that effective differentiated search support, leading to different resources, can now be provided in a digital environment.

4. *Metadata as infrastructure.* Searches are made in and among descriptive metadata. It follows that metadata structures are as much infrastructure as hardware, software, telecommunications, and electrical power.

5. *Restoring the functionality of the library reference collection.* The familiar reference collection of the paper-based library has carefully selected resources well-organized to enable a learner to find out about the context, background, history and relationships of any topic of interest. This functionality is inadequately provided in the digital environment. We believe that this project points to how more adequate provision could be designed.

CONCLUSION

Learning, if it is to be more than memorizing, requires an understanding of context. The Web greatly increases the range of accessible resources capable of providing context. However, internet-accessible resources vary greatly in the metadata vocabularies used to describe and index their resources and in the search support provided. The use of standards should, of course, be encouraged, but mandating the use of the same systems or the same vocabularies is neither possible nor desirable. However, the wider adoption of improved best practices can greatly facilitate the incremental development of interoperability.

Searching by topic requires mapping between different topical vocabularies and automated methods can often provide inexpensive but useful solutions. Place name gazetteers are pivotal for geographic search because they link places with spaces and enable both map visualizations and analysis of spatial relationships between places. A named time period directory can play an analogous role for events, dates, and chronological relationships. Searching for people by name is a well-understood problem, but relating the events in their lives to contextualizing resources needs more development. Searching can be supported by embedding live queries in mark-up and by the use of meaningful codes in structured URLs. Picking a museum object at random and finding what materials of different kinds elsewhere are most closely related to it and to its context requires substantially better indexing and interoperability

than is currently provided, but a series of improvements in standards and best practices will, individually and cumulatively, help us to advance in that direction.

ACKNOWLEDGMENTS

Vincent Briggeman, Kimberly Carl, Sarah Ellinger, Matt Meiske, Linda-Cathryn Muehlinghaus, Vivien Petras, Jeanette Zerneke, and others contributed to the work reported. We also thank our very helpful Advisory Committee members: John Ober, Merrilee Proffitt, Sherry Volk, and Raymond Yee

REFERENCES

A list of more than twenty related papers, presentations, and reports is provided (and will be added to) at <http://ecai.org/imls2004/publications.htm>. The most important item are:

- Bringing lives to light: Biography in context.* (2006). [Project website]. <http://ecai.org/imls2006/>
- Buckland, M. K. (2006). Description and search: Metadata as infrastructure. *Brazilian Journal of Information Science* vol. 0 (2006). Preprint: <http://www.ischool.berkeley.edu/~buckland/Brazil06.pdf>
- Buckland, M. K. (2007a). The digital difference in reference collections. *Journal of Library Administration* (Forthcoming, also separately as proceedings for the *Conference on Printed Resources and Digital Information: The Future of Coexistence*. Oklahoma University Libraries, Oklahoma City. March 2 & 3, 2006. Preprint: <http://www.ischool.berkeley.edu/~buckland/OKLABuckland.pdf>
- Buckland, M. K. (2007b). Naming in the library: Marks, meaning and machines. In *Nominalization, nomination and naming in texts*, ed. by Christian Todenhagen & Wolfgang Thiele. Tübingen, Germany: Stauffenburg Verlag, forthcoming. Preprint: <http://www.ischool.berkeley.edu/~buckland/naminglib.pdf>
- Buckland, M. K., A. Chen, H.-M. Chen, Y. Kim, B. Lam, R. R. Larson, B. Norgard & J. Purat. (1999). Mapping entry vocabulary to unfamiliar metadata vocabularies. *D-Lib Magazine* 5, no. 1 (Jan 1999). <http://www.dlib.org/dlib/january99/buckland/01buckland.html>
- Buckland, M. K., A. Chen, F. C. Gey & R. R. Larson. (2006). Search across different media: Numeric data sets and text files. *Information Technology and Libraries* 25, no 4 (Dec 2006): 181-189. <http://metadata.sims.berkeley.edu/searchacross.pdf>
- Buckland, M. K., A. Chen, F. C. Gey, R. R. Larson, R. Mostern & V. Petras. (2007). Geographic search: Catalogs, gazetteers, and maps. *College & Research Libraries* Forthcoming Sept 2007. Preprint: <http://metadata.sims.berkeley.edu/geographicsearch.pdf>
- Buckland, M. K., F. C. Gey & R. R. Larson. (2002). *Seamless searching of numeric and textual resources. Final report on Institute of Museum and Library Services National Library Leadership Grant No. 178*. Berkeley: University of California, School of Information Management and Systems. <http://metadata.sims.berkeley.edu/papers/SeamlessSearchFinalReport.pdf>
- Buckland, M. K., F. C. Gey & R. R. Larson. (2004). *Going places in the catalog: Improved geographic access: Final report*. http://ecai.org/imls2002/imls2002-final_report.pdf
- Buckland, M. K. & L. R. Lancaster. (2004). Combining time, place, and topic: The Electronic Cultural Atlas Initiative. *D-Lib Magazine* 10, no. 5 (May 2004). Consulted January 28, 2007. Available at <http://www.dlib.org/dlib/may04/buckland/05buckland.html>

- Buckland, M. K. & L. R. Lancaster. (2006). Advances in discovery: The Electronic Cultural Atlas Initiative experience. *First Monday*. 11, no. 8 (August 2006).
http://www.firstmonday.org/issues/issue11_8/buckland/index.html
- Electronic Cultural Atlas Initiative. (2003). *ECAI Iraq*. <http://ecai.org/iraq/>
- Frontiera, P., R. Larson & J. Radke. (2007). A comparison of geometric approaches to assessing spatial similarity for GIR. *International Journal of Geographical Information Science*, (in press).
- Going Places in the Catalog: Improved Geographical Search*. Project website ecai.org/imls2002
- Larson, R. R. (2007). Domain specific retrieval: Back to basics. In *CLEF 2006 Proceedings*. Springer (LNCS in press).
- Petras, V. (2006). *Translating dialects in search: Mapping between specialized languages of discourse and documentary languages*. PhD dissertation, University of California, Berkeley.
- Petras, V., R. R. Larson & M. K. Buckland. (2006). Time period directories: A metadata infrastructure for placing events in temporal and geographic context. In *Opening information horizons: Joint Conference on Digital Libraries (JCDL)*, Chapel Hill, NC, June 11-15, 2006. <http://metadata.sims.berkeley.edu/tpdJCDL06.pdf>
- Petras, V., M. Meiske, R. Larson, J. Zerneck, K. Carl & M. Buckland. (2005). Leveraging Library of Congress Subject Headings to improve Search for Events: A Time Period Directory. *Technical Report*. 2005. <http://metadata.sims.berkeley.edu/tpd/TPD-report.pdf>
- Zerneck, J. L., M. K. Buckland & K. Carl. (2006). Temporally dynamic maps: The Electronic Cultural Atlas Initiative experience. *Human IT* 8.3: 83–94. <http://www.hb.se/bhs/ith/3-8/jzmbkc.pdf>