University of California, Berkeley
Electronic Cultural Atlas Initiative

Bringing Lives to Light: Biography in Context

A National Leadership Project for Libraries funded by the

INSTITUTE OF MUSEUM AND LIBRARY SERVICES
GRANT No. LG-06-06-0037-06

FINAL REPORT FOR THE PERIOD
OCTOBER 1, 2006 THROUGH SEPTEMBER 30, 2009

December 15, 2009, rev. February 16, 2010

The project was officially funded effective October 1, 2006. In July, 2008 the project was granted a one year no-cost extension until September 30, 2009. This is our final narrative report on the project, covering the entire period and including a bibliography of papers and technical reports produced for the project.

Project Context

Cultural heritage, history, and social sciences are fundamentally about human activity. Everyone is interested in what other people do and have done.

Life-stories are hard to beat as a basis for narrative and for engaging interest, especially among young people. Biographies are regularly among the best-sellers. Not only History, but also Geography and most other subjects can come alive in the travelogues, journeys of discovery, and the life-stories of those involved. Science can be explained through the work of scientists. Engineering is routinely explained through the heroic struggles of inventors. Even natural history is often taught through the unfolding drama of the activities of an individual animal during its life-cycle or through the seasons of the year.

But mere narrative is not enough. Understanding the context differentiates education from memorizing. Building and supporting a community of learners needs more than facts. It is understanding the circumstances of people's actions that illuminates their lives, but there is a significant gap in the infrastructure developed by libraries, museums, and publishers in this area. We have standards for handling people's names, but not for their lives. There is, quite simply, no established standard or "best practices" for encoding what people do, nor for helping to search out the resources that can provide the context to understand their actions and experiences.

Our objective is to design, demonstrate, and evaluate techniques that would bring lives to light by revealing them in their contexts.

Approach to Biographical Information

This project builds on work accomplished in two earlier projects, *Going Places in the Catalog: Improved Geographic Access* (http://ecai.org/imls2002) and *Support For The Learner: What, Where, When And Who* (http://ecai.org/imls2004). The current project is using the metadata infrastructure developed in those projects, including expanded historical gazetteer information and the "Time Period Directory", a novel structure that allows named time periods (such as reigns of particular monarchs, or dynasties) to be associated with their appropriate dates and contextualizing geographic information using gazetteers. The tools developed in these previous projects include dynamic search linkages to online resources such as the online catalogs of the Library of Congress and the University of California as well as many other institutions including Museum and educational resources (See http://www.ecai.org/imls2004/imls4w/).

The remainder of this report will discuss the progress and developments made in each of the tasks outlined in our proposal.

**Task 1: Encoding Biographical Content**

Lives and Events

We have adopted the approach of using RDF as the base structure for our metadata and markup, and have been incorporating the best features from a variety of different metadata schemas in a Semantic Web – compatible structure. A worked example of this approach is available and shows some of the life events of Davy Crockett and Andrew Jackson expressed as RDF/XML: http://people.ischool.berkeley.edu/~ryanshaw/bio/congress/   This example draws from five vocabularies (in addition to RDF, RDF Schema, and Dublin Core):

1) FOAF provides classes and properties for describing people. http://xmlns.com/foaf/spec/
2) BIO provides classes for events, and properties for relating people to events and events to places and times. http://vocab.org/bio/0.1/
3) OWL-Time provides properties for expressing temporal relations among events (so that when specific dates are unknown we can express that this happened after that). http://www.w3.org/TR/owl-time/
4) GeoNames provides classes and properties for describing places: http://www.geonames.org/ontology/
5) Basic Geo (WGS84) provides lat/long properties: http://www.w3.org/2003/01/geo/

Congressional Biography Data

The same markup described above has been used to convert data from the Biographical Directory of the United States Congress (http://bioguide.congress.gov/), combined with other resources including data from Geonames, Dbpedia, and Freebase. We are continuing to link in other biographical texts and text contents from the books scanned by the Open Content Alliance (OCA). This data has been made available using a simple search interface (available at http://

metadata.berkeley.edu/Congress/) and also as Google Earth browser displays using a geographic search-based network link (the KML for the network link is available at http://metadata.berkeley.edu/Congress/Congress.kml - which can be downloaded to run on any machine with Google Earth installed. In the Google Earth version we include place marks for the birthplace and (when deceased) death location for each of the members of Congress from the first congress in 1789 until the present. The screen capture below shows the Google Earth interface for the Congressional Biography data, with the data for John Adams (Washington's Vice President and thus president of Senate, and later 2nd President of the U.S.). When the user clicks on a birthplace "b" icon or death place "d" icon the summary biographical events page with birth, death, and election events is shown, along with links to further contextual information on the person ranging from a link to the official congressional biography, to Wikipedia articles and books about or by the person. The lines shown on the map link the birth and death locations for each person where both events are known. The time bar on the upper left of the display allows the user to select a range of dates and only persons alive during that period will be displayed. We are continuing to expand the contextual information linked to the Congressional Biography data
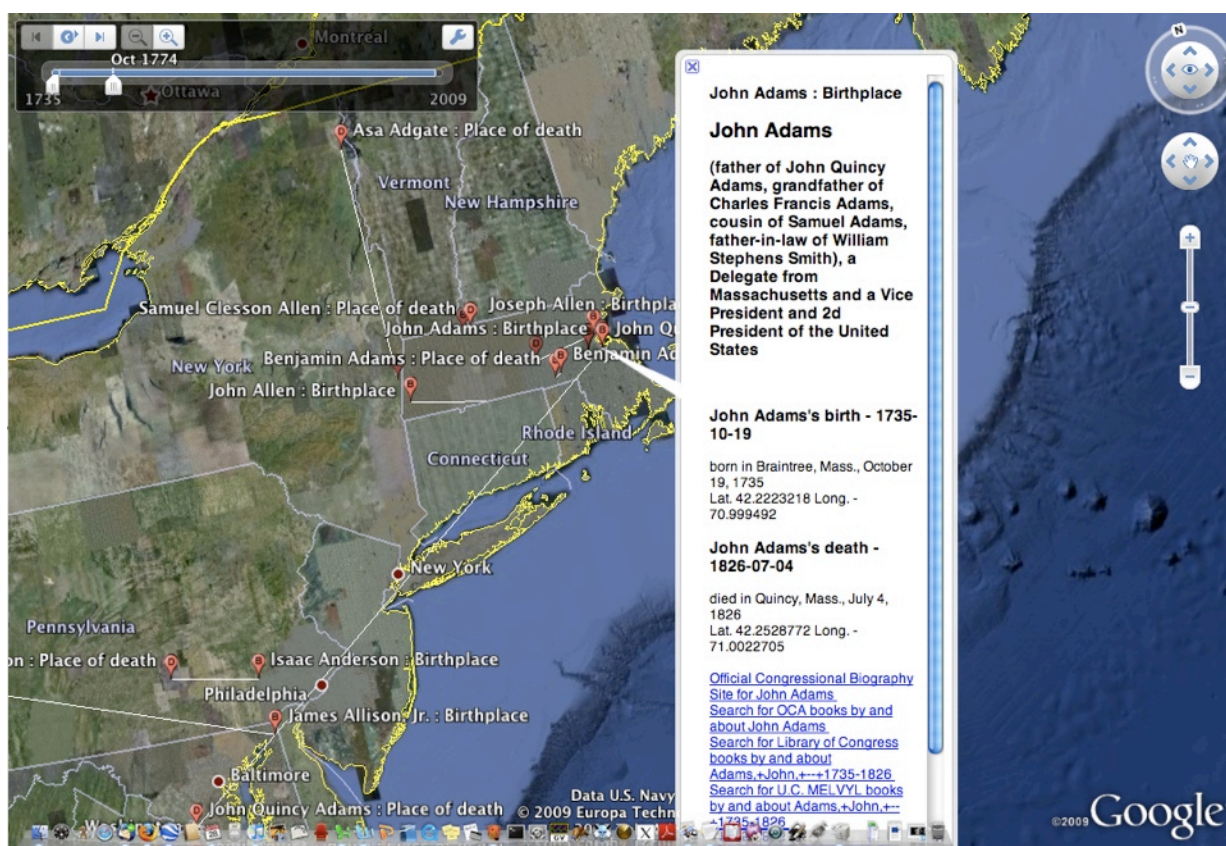
We also plan to create display mechanisms like those developed for Emma Goldman's lecture itinerary and our interface for scanned texts from books or web pages (see http://metadata.berkeley.edu/demos)

To create the Emma Goldman itinerary interface, we began with Barry Pateman's (our collaborating scholar from the Goldman Papers Project) WordPerfect file containing the Emma chronology. We then converted this to the OpenDocument XML[1] format] using Open Office. Because Barry used consistent styling (bold, italics, etc.) and formatting for things like city names and dates in his chronology, we could easily parse this XML and produce an Atom XML [2] document with an entry for each event in the chronology. Then we georeferenced this Atom document using Geonames[3]. Finally, we transformed this Atom file into RDF/XML using classes and properties from the Event ontology[4]. We also transformed it into KML[5] for viewing in Google Earth like the Congressional data shown above.

The prototype consists of a single XHTML file with the RDF triples encoded as RDFa[6] and embedded in HTML attributes. Using the open source Yahoo User Interface Javascript libraries[7], Google Maps Javascript libraries[8], and our own Javascript code, we created the interactive browsing and searching functionality which searches within the HTML document for matching RDF triples upon selection of a place and/or date range.

When a user clicks on the "related document" links on the right-hand event list a popup window (shown above) shows geographically and time and topic-related news stories from regional newspapers, contemporary photos, posters, or other contextualizing documents.

[1] http://en.wikipedia.org/wiki/OpenDocument
[2] http://en.wikipedia.org/wiki/Atom_(standard)
[3] http://en.wikipedia.org/wiki/GeoNames
[4] http://motools.sf.net/event/event.html
[5] http://en.wikipedia.org/wiki/Keyhole_Markup_Language
[6] http://en.wikipedia.org/wiki/RDFa
[7] http://en.wikipedia.org/wiki/Yahoo!_UI_Library
[8] http://code.google.com/apis/maps/

The Irish Interface

A similar concern with providing access to biographical descriptions influenced our interfaces for Irish texts in the related project entitled *Contexts and Relationships: Ireland and Irish Studies* http://ecai.org/neh2007

For the scanned page images, we transformed the TIFF[1] page images and XML OCR output we received from our collaborators at the Centre for Data Digitisation and Analysis at the Queen's University, Belfast, into PNG[2] images and JSON[3] for fast, lightweight viewing in a web browser. The scrolling reading interface is built with Javascript, again utilizing the Yahoo User Interface libraries. Because the OCR output contains information about the bounding boxes

of the recognized words, we can simulate highlighting and selection of words on the scanned image. The named entity detection is done using OpenCalais[4].

[1] http://en.wikipedia.org/wiki/Tagged_Image_File_Format
[2] http://en.wikipedia.org/wiki/Portable_Network_Graphics
[3] http://en.wikipedia.org/wiki/JSON
[4] http://www.opencalais.com/

As noted in previous reports we have put up a visualization generated directly from the example RDF markup for biographical events, using the Exhibit tool from MIT's Simile project. During this reporting period we have been using an extended version of the Simile tool developed by one of the Master's students in the School of Information (Nick Rabinowicz). We are using this tool (called TimeMap) as another display technique for the congressional data described above. The following image is the result of a search displayed using the TimeMap software. The bars along the top are a timeline for particular persons, and places of birth for those persons are shown on the map. Both the timeline and the map can be scrolled, and the icons on the map selected for information similar to that shown in the Google Earth example above.
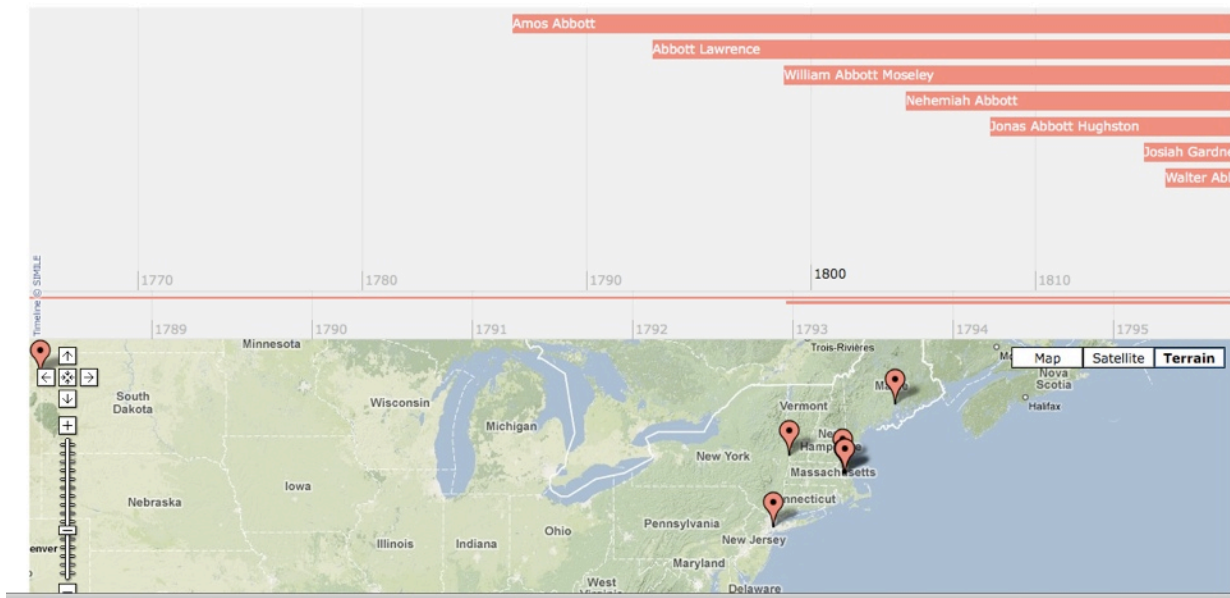
Meanwhile we have been continuing developing methods for extracting biographical event information from texts. In the process we have become aware of additional resources that have recently become available, including the work of DBPedia. According to its developers

"DBpedia is a community effort to extract structured information from Wikipedia and to make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia and to link other datasets on the Web to Wikipedia data." Amongst the data extracted by the DBPedia effort are basic facts from Wikipedia about more than 80,000 people. All of the DBPedia data is available for access from: http://dbpedia.org/ in RDF form that can be incorporated with our developing schema. During this period we have been incorporating supporting references from DBPedia in our prototypes.

Note On Event-Based Mark-Up

During the period being reported, the feasibility of using events as an analytical device for biographical texts was further explored. The idea is that lives are largely composed of a series of events and activities, some brief, some lengthy: birth, education, marriage, military service, etc. In each case there is a kind of action or activity in one or more places during some period of time and commonly involving other people, which lends itself to a 4W (What, Where, When, and Who(else)) structure. The more completely such a 4W framework can be developed, the more effectively the person's life can be contextualized within explanatory resources.

A preliminary study examining in detail numerous brief biographical accounts of several individuals confirms that this approach is feasible and that initial manual analysis is likely to be very useful in generating training data for subsequent computer-assisted analysis. However, there are a number of difficulties: Often the when and where are unspecified or only implied; named institutions need to be included; and one needs to accommodate the tendency for events to be defined in relation to other events. Some of this work was reported on at the ACM/IEEE Joint Conference on Digital Libraries (JCDL) and at the European Conference for Digital Libaries (ECDL) for both 2008 and 2009.

Automatic Markup of Textual Materials

We have been investigating and applying various natural language processing techniques for automatically discovering person and place names in text and marking them up for later processing and display.   In particular we have utilized the Golden Gate XML editor from University of Karlsruhe in Germany as well as the ClearForest (now OpenCalais) analytical text technology in preparation of our demonstration prototypes.   When applied to the Emma Goldman travel itinerary narrative, 478 personal names were automatically identified, as well as 491 cities and towns to which Goldman traveled. A 'venue' recognizer was written to automatically detect where Goldman lectured, 519 places such as "Knights of Pythias Hall" or the "Hod Carriers Hall" were automatically detected and marked up.   A summary of this was written up in the paper presented at the "Information Access to Cultural Heritage" workshop in Aarhus, Denmark in September 2008.

LODE: Linking Open Descriptions of Events

People conventionally refer to an action or occurrence taking place at a certain time at a specific location as an event. This notion is potentially useful for connecting individual facts recorded in the rapidly growing collection of linked data sets and for discovering more complex relationships between data. Doctoral student Ryan Shaw, along with collaborators from

EURECOM in France and the Centrum Wiskunde & Informatica (CWI) in the Netherlands have developed LODE (Linking Open Descriptions of Events) that provides an overview and comparison of existing RDFS+OWL event models, looking at the different choices they make of how to represent events based on earlier work with this project. LODE describes a recommended model for publishing records of events as Linked Data. In the technical report on LODE linked from our publications web site (http://metadata.berkeley.edu/imls2006publ.htm), the authors present tools for populating this model and a prototype of an "event directory" web service, which can be used to locate stable URIs for events that have occurred and to provide RDFS +OWL descriptions of them and links to related resources. The event directory is populated with event information from our Congressional Biographies and Goldman Papers work as well as other information derived from Wikipedia. The LODE interface is shown below with data from the Congressional Biographies.

Events with people and places using Congressional Biography Data

Selecting a place limits displayed events to those occurring at that place

The Biography Light Ontology

The work described above, and much more "behind the scenes" analysis on people and the events in their lives led to a continuous development and revision of various approaches to biographical markup for texts. One of the final products of this work has been the development of the "Biography Light Ontology" which provides a conceptual framework for the events in people's lives and a representation method using semantic web approaches. The Ontology and its testing on biographical data sources from our partners in this project, The Emma Goldman Papers and the U.K. Archives Hub, is described in detail in the Technical report by Michele R. Ramos (http://metadata.berkeley.edu/imls2006publ.htm). The Biography Light Ontology is based on the ontology developed for the LODE project above. These ontologies are depicted below as Figure 1 and 2 from the report.
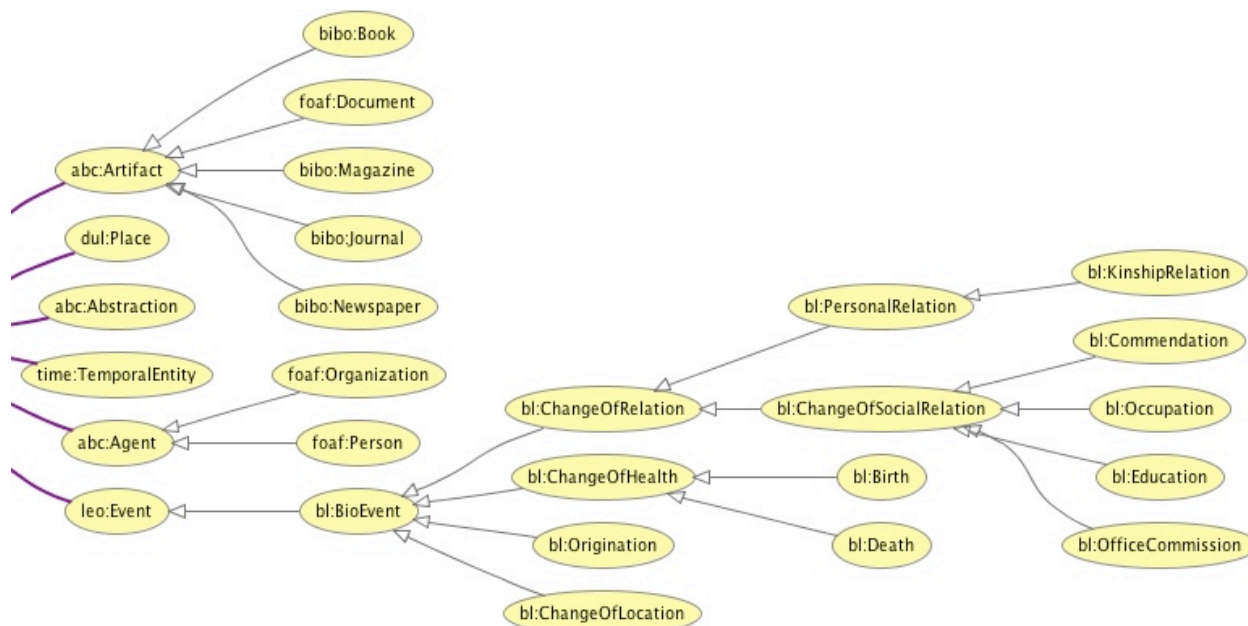
Figure 1: Image of the Biography Light model integrated with LODE, ABC, Bibo, and FOAF.
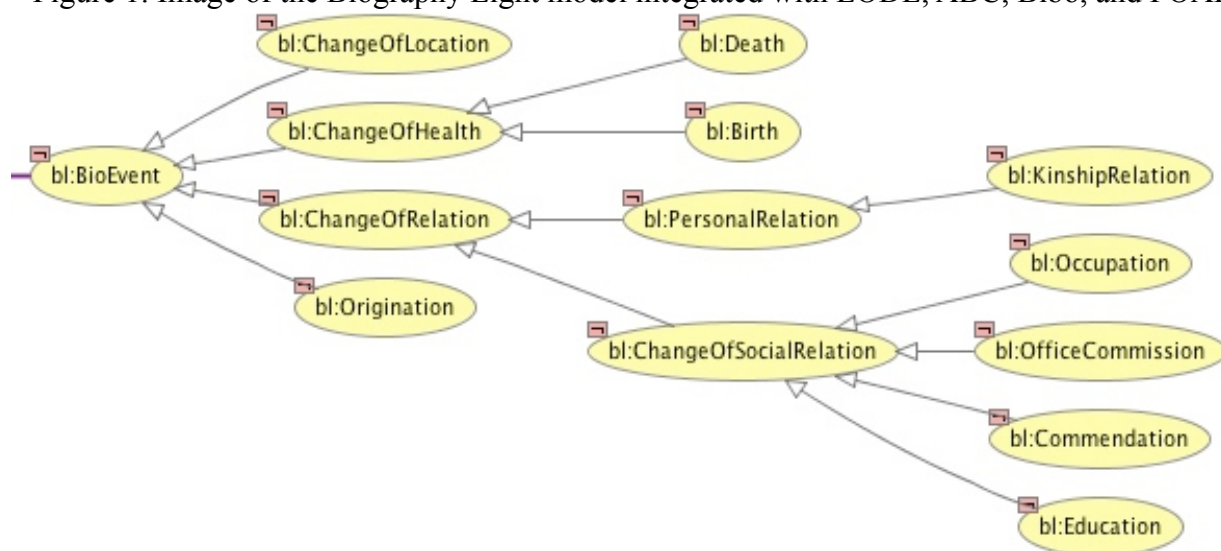
Figure 2: The Biography Light Ontology

The Biography Light class definitions for biographical events include the following classes and properties:

CLASSES

Name: BioEvent
Subclass of: leo:Event
Equivalent to: bio:Event
Description: A class representing biographic events associated with a person. This class does not differentiate between changes of states or situations, actions, or processes. It is a class

representing a temporal and existential context associated with a person. This class should not be considered as a sub-event of a larger event instance, rather the class of biographic events should be considered as a type or kind of occurrence.

Name: ChangeOfHealth
Subclass of: bl:BioEvent
Description: A class representing events related to the health of a person.

Name: Birth
Subclass of: bl:ChangeOfHealth
Equivalent to: bio:Birth
Description: A class representing the event of a birth of a person.

Name: Death
Subclass of: bl:ChangeOfHealth
Equivalent to: bio:Death
Description: A class representing the event of a death of a person.

Name: ChangeOfLocation
Subclass of: bl:BioEvent
Description: A class representing biographic events involving the change of location for a person. Events of this type include (but are not limited to) traveling, migration, or immigration and can be considered as involving a change of state or situation.

Name: Origination
Subclass of: bl:BioEvent
Description: A class representing the creations, discoveries, or inventions produced by a person. Events of this type can be considered an activity or process involving the person, were the result is the initiation of some object (such as an artifact or document) or an abstraction (such as a concept or theory).

Name: ChangeOfRelation
Subclass of: bl:BioEvent
Description: A class representing a general biographic relationship between agents, such as people, families, corporate bodies, organizations, or institutions.

Name: PersonalRelation
Subclass of: bl:ChangeOfRelation
Description: A class representing general biographic relationships between two people. Events of this type can be considered a situation, role, or status that may not have a clear beginning or end date, nor necessarily take place at a specific location. Examples of events of this type include (but are not limited to) friend, colleague, mentor, apprentice, or co-habitants.

Name: KinshipRelation
Subclass of: bl:PersonalRelation
Description: A class representing kinship relationships between two people. This class includes (but is not limited to) legal (i.e. guardian), honorific (i.e. godparents), affinal (i.e. marriage), and consanguineal (i.e. descent) kinship relationships.

Name: ChangeOfSocialRelation
Subclass of: bl:ChangeOfRelation
Description: A class representing general biographic relationships between a person and larger social institutions or groups, such as organizations, groups, or political, religious, and social communities. Events of this type can be considered a situation, status, or social relationship that may not have a clear beginning or end date, nor necessarily take place at a specific location.

Name: Commendation
Subclass of: bl:ChangeOfSocialRelation
Description: A class of biographic events for awards or prestigious recognitions attributed to a person.

Name: Education
Subclass of: bl:ChangeOfSocialRelation
Description: A class of biographic events pertaining to the education, study, or training of a person.

Name: Occupation
Subclass of: bl:ChangeOfSocialRelation
Description: A class of biographic events pertaining to occupations, professional work, or vocation of a person.

Name: OfficeCommission
Subclass of: bl:ChangeOfSocialRelation
Description: A class of biographic events pertaining to government or religious offices held by a person.

PROPERTIES

Name: relatedEvent
Domain: leo:Event
Range: leo:Event
Description: An object property that relates an instance of an event to another event (or event type). Within biographic texts many dispersed and formally unrelated events may hold relationships to each other, such as a crime event that precedes the event of the trail, or occurrences that may be connected to a larger situation or movement. Sub-event relationships

restrict modeling events to part-whole relationships when events as they are described in biographic texts may not hold this same association.

As an example, we quote from the report by Michele Ramos:

Modeling Experiments and Examples

This model was developed using six biographic excerpts from the UK Archives Hub[1] and six biographies from the Emma Goldman: A Documentary History of the American Years, volume III Light and Shadows (1910-1916) (see Appendices). The process of encoding biographic events, involved translating the texts into factoids of a basic *4W* structure: *What* was the event, *Who* was involved, *When* did the event occur, and *Where* did the event take place. Below is an

Robert George Collier Proctor	Proctor, Robert George Collier Proctor
(1868-1903), bibliographer, was born in Budleigh Salterton, Devon, on 13 May 1868. He was educated at a preparatory school in Reading and at Marlborough College, before joining Bath College in 1881. In 1886 he entered Corpus Christi College, Oxford (B.A., 1890). Proctor worked at the Bodleian Library from 1891 to 1893, where he catalogued incunabula. In 1893 he became assistant in the printed books department at the British Museum, where he was employed for the rest of his life. Proctor devised a new Greek type and compiled an Index of early printed books to the year MD (1898). He disappeared while on a walking tour in Tyrol in 1903.	ChangeOfHealth ◆birth, 1868-05-13, Budleigh Salterton, Devon ◆death, 1903, disappeared in Tyrol (?) ChangeOfSocialRelation ◆studied at Marlborough College, before 1881 ◆studied at Bath College in 1881 ◆studied at Corpus Christi College, Oxford, in 1886-1890 ◆employed by Bodleian Library, between 1891-1893 ◆employed by the British Museum in 1893-1903 Originatation ◆discovered a new Greek Type ◆authored work "Index of early printed books to the year MD" (1898)

example biography with the corresponding factoid outline.

Table 1: Example biography and corresponding event factoids.

The right side of the table delineates the event factoids that will be modeled as instances of corresponding biographic event classes. The process of decomposing biographic texts into these event statements requires emphasizing certain elements while excluding others. For example, Proctor's education at Corpus Christi College is modeled over an interval of time, as opposed to two separate events such as his entering college and eventual graduation. The details of his employment are not specified nor are his levels of education. What is left is a framework

---

[1] UK Archives Hub, http://www.archiveshub.ac.uk/ one of our collaborators in the Bringing Lives to Light Project.

that connects Proctor to biographic events, organizations, objects, places, and dates. Visualizations of this example are provided in Figure 4.

This methodology of modeling biographic event factoids does have a series of problematic aspects. One is the process of translation itself. Translating biographic texts into factoid statements makes automatic parsing and event extraction difficult at best. Translation also brings up the important issue of a controlled vocabulary for event statements. Is it better to translate biographic information generally, stating simply that Person A worked with Organization B and excluding the details of their role within the organization, or is it better to keep the factoid statements as close to the original description as possible? There is also the issue of modeling other biographic subjects within a biographic text. Additional information about other biographic subjects, such as birth and death dates, is quite common among biographic texts. Consider this excerpt from the short biography of Leonard Dalton Abbott:

> Abbott was associate editor of *Current Literature* (1905-1925), edited *Free Comrade* (1912) with J. William Lloyd and C. L. Swartz, and *Modern School* in 1912 with Harry Kelly. Written works include: *Francisco Ferrer: His Life, Work, and Martyrdom* (ed.; Francisco Ferrer Association, 1910), *What Ferrer Taught in His Schools* (Current Literature Pub. Co., 1910), and *The Detroit Francisco Ferrer Modern Schoo*l (with William Thurston Brown; New York: Herold Press, 1912).
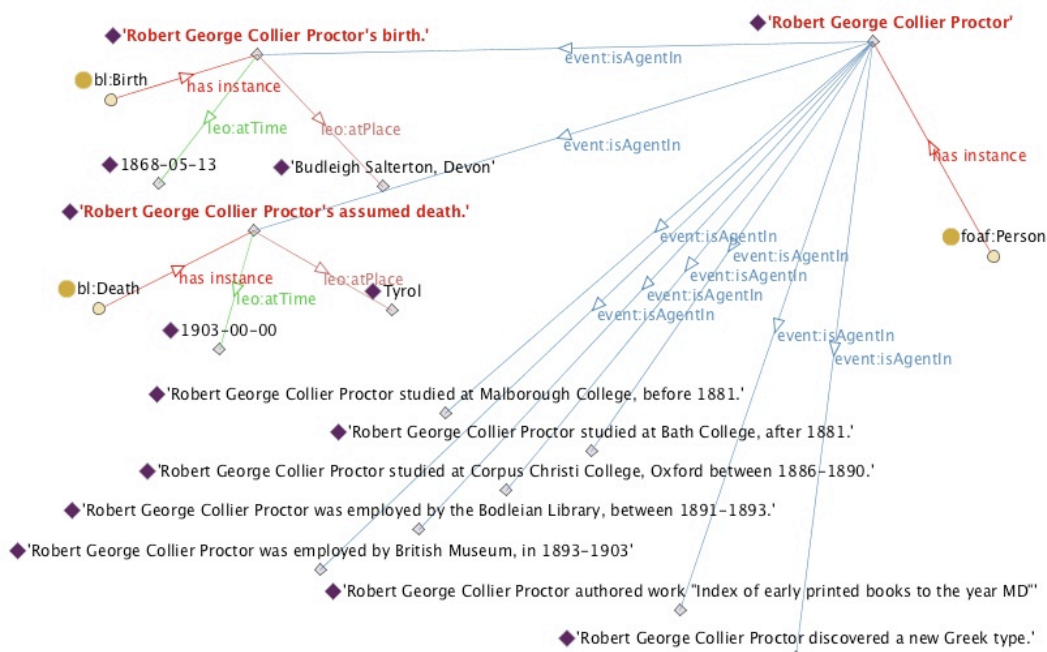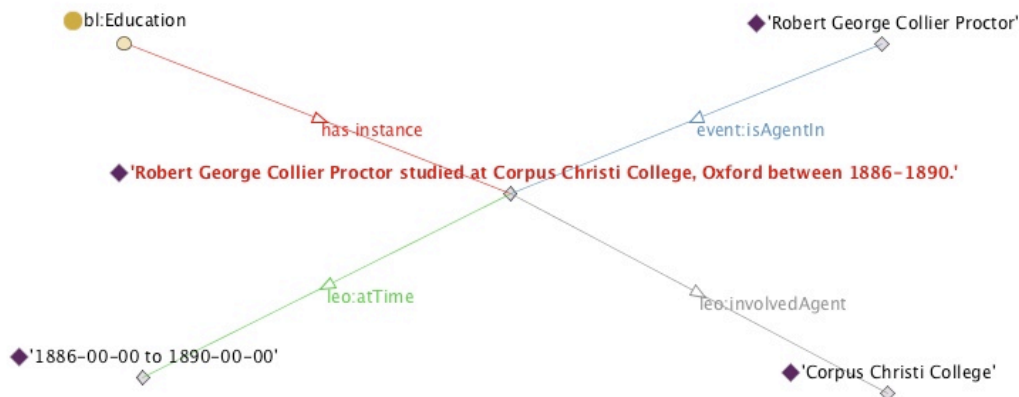
Figure 4: Image and example of Proctor's biographic events.

There are many different approaches one could take in modeling Abbott's created works in relation with other people. We could model the event of Abbott authoring or editing a work as involving another person, or we could independently model the information stating separately that Harry Kelly, J. William Lloyd, and C.L. Swartz were editors of the corresponding magazines and linking them to the primary biographic subject through a relationship instance, for example "Leonard Dalton Abbott was colleague of Harry Kelly." When it comes to modeling information, there is no right or wrong way, only good or bad modeling practices. Choice depends on the purpose. It is up to the translator or encoder to include inferred information, such as time periods or relationships, or to exclude events that do not establish basic links between people, places, events, dates, and objects or concepts. It is also up to the translator or encoder regarding specificity or generality when translating biographic events into factoids. The purpose of this model is to provide a framework for common biographic events as a supplement or addition to the original text – not a substitution.

While we have made significant progress in the identification of events and their automatic markup within texts, as the above examples show, there is still a ways to go. In the following section we examine the how identified information in texts can be used to connect to the broader context of events through embedded links and queries.

Task 2: Interoperability through embedded links and queries

Throughout the project we have been expanding the linkages and scope of the tools for providing access to biographical resources. As the screen capture shows (Figure 5), the various places and persons identified by the analysis of the scanned text are provided with links to a variety of related and supporting information resources. The image shows the links for the "Ask about Ireland" web site, The Irish Resources in the Humanities site, the CELT corpus of electronic texts, COPAC, Google Books, Google Scholar and Wikipedia. The same resources can be searched for any named place or person identified in the scanned text.

The "Context Finder" shown in that image is also an aspect of the Congressional search Google Earth Interface shown in the subsequent set of images Figures 6-11 (from a talk presented at the ECDL Books Online Workshop in Corfu). A person serving in Congress from the first Congress in 1789 to the present is included in the RDF-based dataset. The Google Earth

interface allows the user to browse by locations of Birth and Death, as well as allowing the results to be restricted to particular time periods (using the time bar shown in Figure 6) and for any birth or death event, a location marker can be selected to show the complete metadata for a given person (as described above in the section on the Congressional Biography data) and dynamically generated search links for a variety of resources. Figure 7 shows the result of selecting the Official Congressional Biography link, and Figure 8 shows the result of selecting the Wikipedia link. Figures 9-11 show the John Adams information, and how selecting the OCA Books link leads to a catalog search (in a collection of MARC developed for a subset of OCA Books used in the INEX Books track) with results that include links to the actual OCA scanned books hosted at the Internet Archive.

Figure 5. Context Finder Interface for Scanned Books
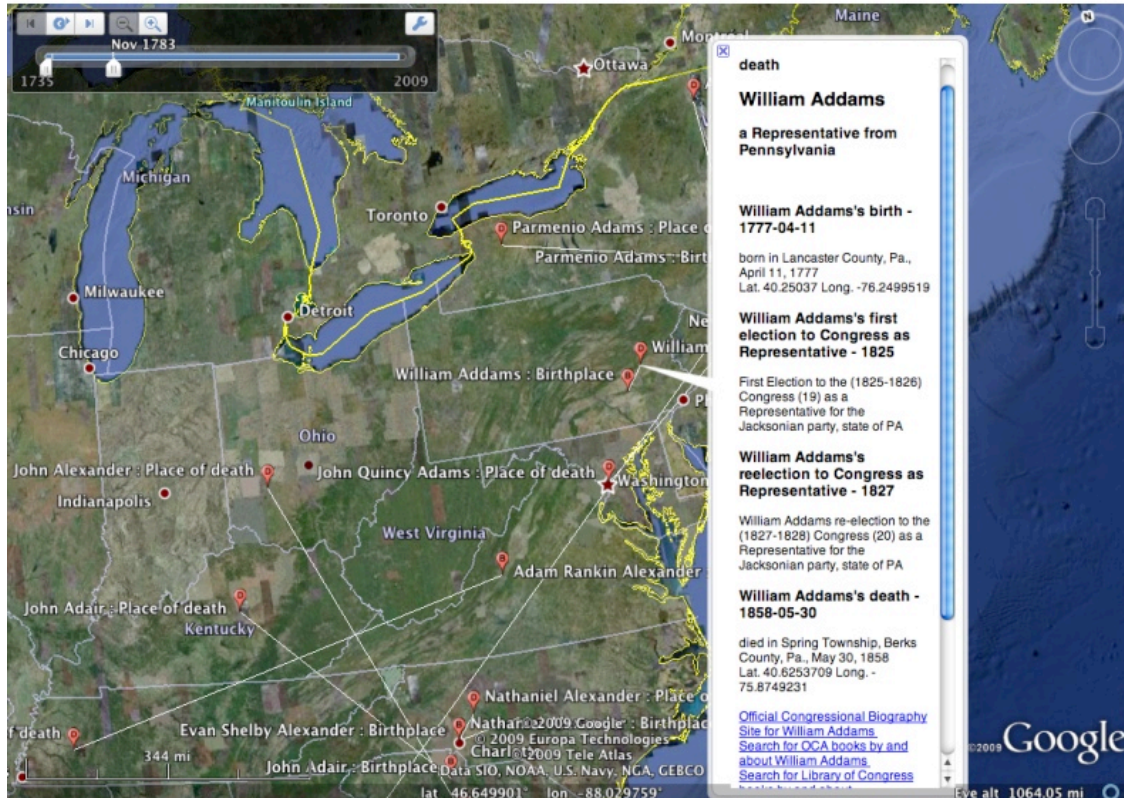
Figure 6: Congressional GeoTime Searching with Context Linking
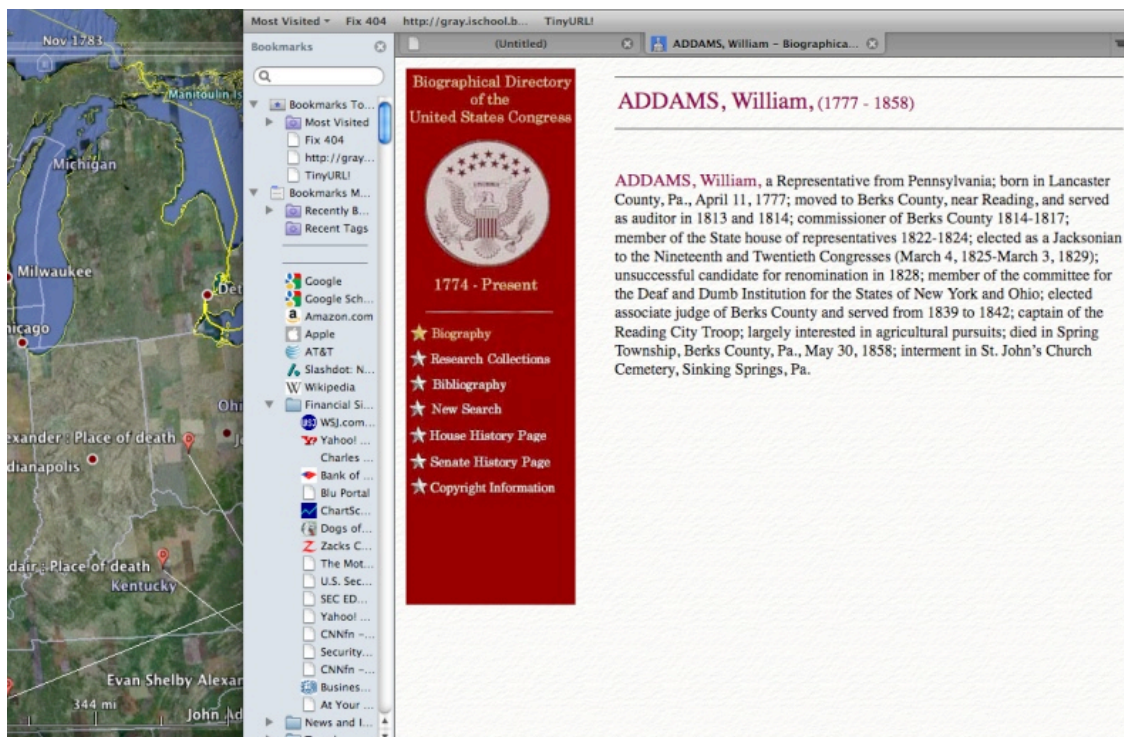
Figure 7. Linking to the Official Congressional Biography from the Map Entry
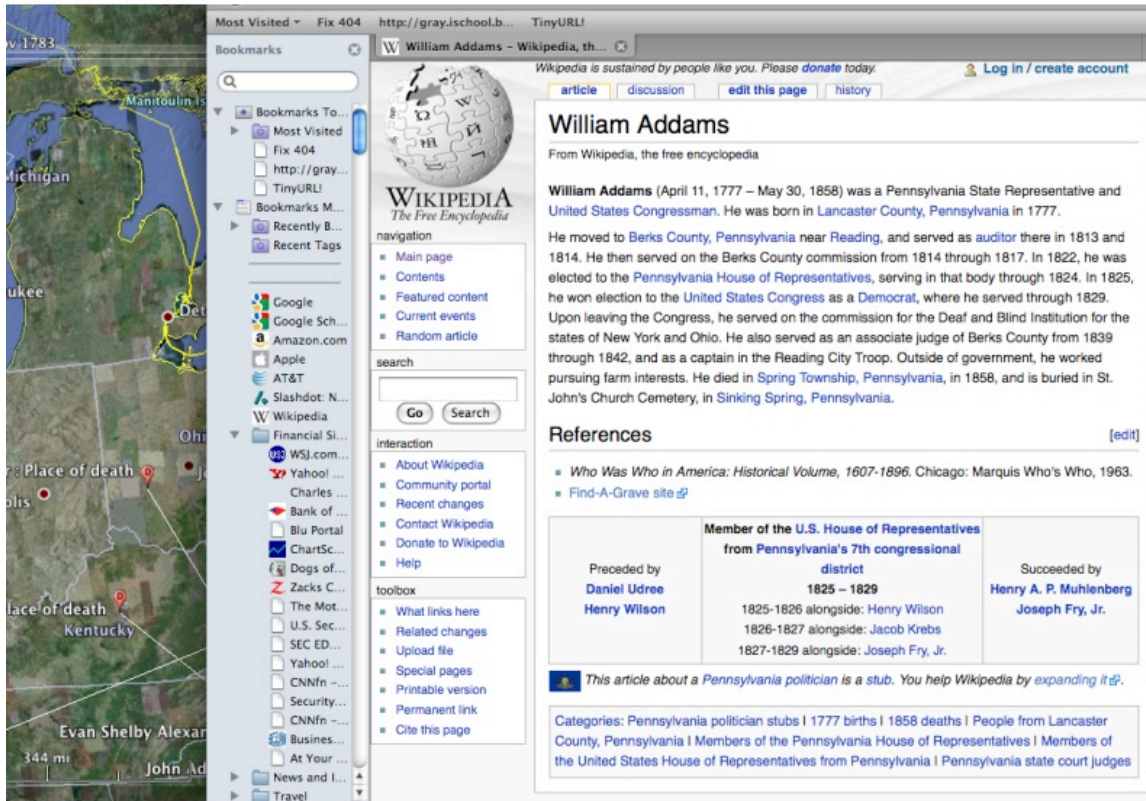
Figure 8. Linking to Wikipedia from Map Entry.
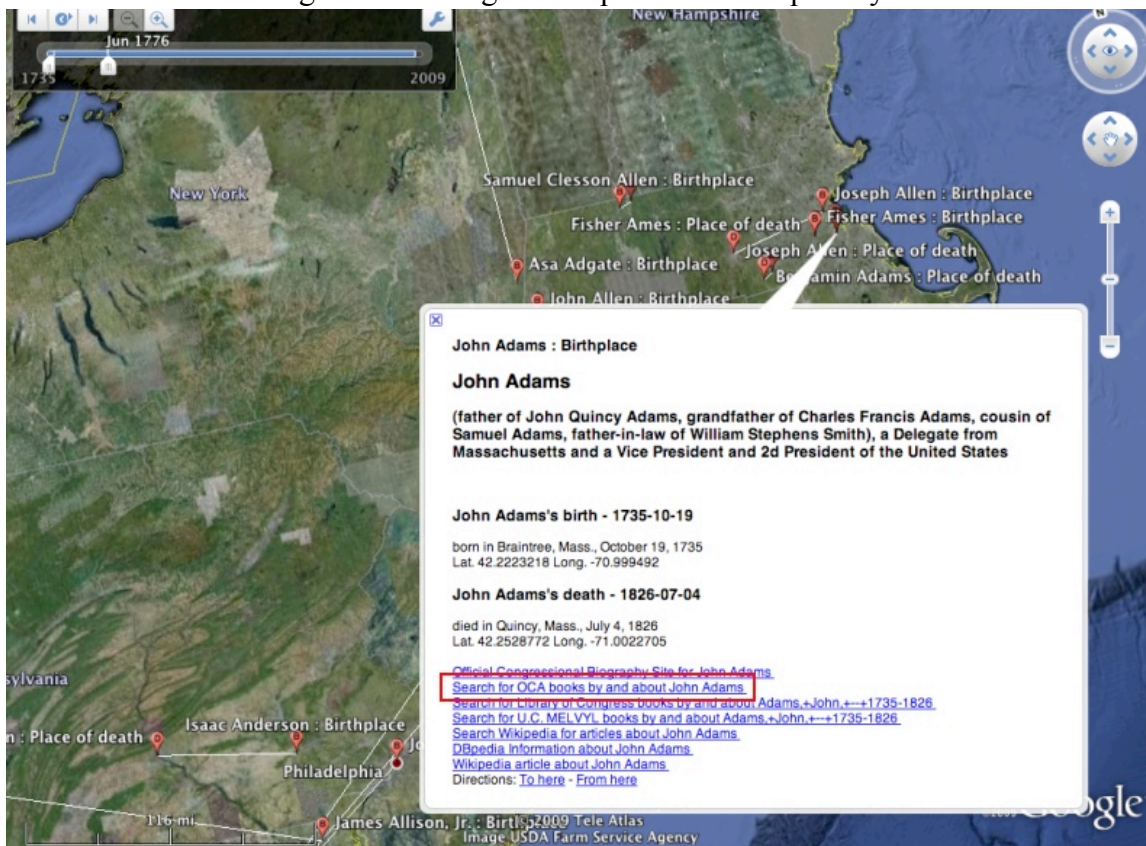
Figure 9. Congressional GeoTime Searching For Open Content Alliance Books.

Figure 10. Open Content Alliance Book search Result.
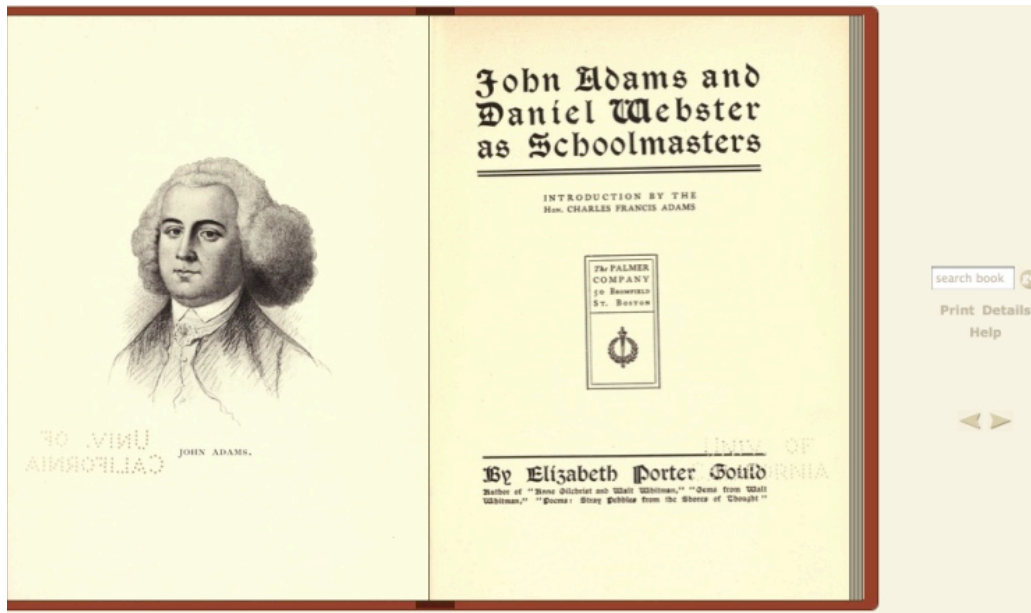
Figure 11: OCA book viewer (via Internet Archive). Could also be loaded to the Context Finder Reader.

Task 3: Tools for Effort-Effective Editing

Working with our collaborator Barry Pateman, an editor from the Emma Goldman Papers project, we are investigating editing and markup tools to see what features might be beneficial for his work. As noted above, we have been working with the the GoldenGate editing software which can be used to automatically parse and help suggest appropriate markup for relevant parts of the texts. As the split screen shots below show, the editor can take texts like those on web sites and elsewhere and identify state and county names (among other things configurable in the editor itself. We were not able to integrate the same features into the preferred editing systems of the Goldman Papers editors, but this is being considered for future projects.

Figure 12: Screen shots of OpenCalais editor and Congressional biography.

Task 4: Testing in Practical Applications

Work developed for this project (including the Biography Light ontology and the LODE system, as well as the Google Earth interfaces for Congress and the Emma Goldman Lectures site are working versions available for use. In addition, technology developed for this project is being used extensively in the joint IMLS-NEH-funded project "Context and Relationships: Ireland and Irish Studies" (some of the above examples show the "Irish" connection). In addition, work developed during this project is an integral part of the technology being proposed for future projects jointly with the University of Virginia Institute for Advanced Technology in the Humanities (IATH), specifically for extracting biographical information from EAD records (as shown in the Biography Light example above) and using this information to help populate

Encoded Archival Context: Companies, Persons and Families (EAC-CPF) records for use and sharing in the archival community. Berkeley's tasks in that project will be the adaptation and continued development of the software developed in this project to efficiently and accurately derive EAC-CPF records from existing archival finding aids; and enhancing them by matching them against records in the Library of Congress Name Authority File (LCNAF) and the Getty Vocabulary Program Union Lists of Artist's Names (ULAN). The software will be developed and tested using EAD-encoded finding aids from the Library of Congress (LoC) and three consortia, the Online Archive of California (OAC), the Northwest Digital Archive (NWDA), and Virginia Heritage (VH), as well as other archival collections (as new partners are added).

Task 5: Documentation and Dissemination

We have been publishing work based on this project in a number of conference and journal venues. The papers below include papers and presentations that we have identified as related to this project. This list is also not complete, since we are continuing to present and publish work related to this project in other projects. See list at http://metadata.berkeley.edu/imls2006publ.htm

M. Buckland & R. Larson. "Infrastructure for the representation of events and their contexts." Poster presentation, *American Society for Information Science and Technology, Austin, TX*, Nov 7, 2006.

Michael Buckland. "Incorporating Historical and Geographical Discussions into a Search Interface." *Association of American Geographers Annual Meeting, San Francisco*, April 17. 2007.

Michael Buckland. "The reference collection in the digital library." ECAI Congress of Cultural Atlases III: Time & Space in Eurasia, held in conjunction with the 7th International Conference of the Historical Sources of Eurasian and North African Civilisations: Computer Approaches, Moscow, Russia, May 29-June 1, 2007. May 30, 2007.

Michael Buckland. "The Electronic Cultural Atlas Initiative: 10 years experience with geo-temporal resources." International Symposium on Digital Earth June 7, 2007.

Michael Buckland. "Geography, Time, and the Representation of Cultural Change -- Experience from a Large Collaboration: The Electronic Cultural Atlas Initiative." Knowbynet conference, organized by the Institute for Museum Research, State Museums of Berlin / Foundation Prussian Heritage, Berlin, Germany. June 22, 2007. http://people.ischool.berkeley.edu/~buckland/berlin07.pdf http://www.knowbynet.de/download/contibutions/buckland_kbn.ppt

Ray R. Larson. "The Future of Search." Presentation at the California Association of Research Libraries (CARL), San Francisco, September 7, 2007.

Ray R. Larson. "Bringing Lives To Light: Browsing and Searching Biographical Information with a Metadata Infrastructure" (Demonstration & Poster) European Conference on Digital Libraries, Budapest, Hungary, September 18, 2007.

Michael Buckland. "Reference Library Service in a Digital Environment: A Question; an Explanation; and a Solution." Friday Afternoon Seminar, School of Information, UC Berkeley, Sept 21, 2007.

Michael Buckland, Aitao Chen, Fredric C. Gey, Ray R. Larson, Ruth Mostern, and Vivien Petras. "Geographic Search: Catalogs, Gazetteers, and Maps." *College and Research Libraries*, Vol 68, No. 5 (September 2007), pp. 376-387.

Michael Buckland. "Linking texts to external resources for context and relationships." Electronic Cultural Atlas Initiative / Pacific Neighborhood Consortium Joint Conference, Berkeley, Oct 18-20, 2007. Oct 19, 2007.

Fredric Gey. "Multi-genre Search Using Common Geography". Presented at the Biennial Meeting of the Interuniversity Consortium for Political and Social Research at the University of Michigan, October 20, 2007.

Ray Larson, Michael Buckland and others. Workshop on biographical mark-up. Electronic Cultural Atlas Initiative / Pacific Neighborhood Consortium Joint Conference, Berkeley, Oct 18-20, 2007. Oct 20, 2007.

Michael Buckland. "The '9 to 5' Problem; or, Reference Library Service in the Digital Library Environment is Back-to-Front, Upside-down, and Inside-out – and What to do About it." American Society for Information Science and Technology Annual Meeting, Milwaukee, Oct 20-24, 2007. SIG HFIS session: Revisiting the Foundations of Information Discovery and Access Systems, Oct. 21, 2007

Ray Larson & Michael Buckland. "Browsing with a metadata infrastructure for events, periods, and time." [Poster]. American Society for Information Science and Technology Annual Meeting, Milwaukee, Oct 20-24, 2007. Oct 23, 2007.

Michael Buckland, Kim Carl, Barry Pateman & Ryan Shaw. "Self-service reference." Academic Library 2.0, a conference sponsored by the Librarians Association of the University of California, Berkeley Division, on November 2, 2007, Berkeley.

Ray R. Larson. "The Electronic Cultural Atlas Initiative (ECAI): Time and Place for Cultural Access. Invited presentation for the EU DigMap Initiative, Lisbon, Portugal, December 7, 2007.

Ray R. Larson. "Geographic Information Retrieval and Results Display". Invited presentation, Computer Science Department, University of Duisberg/Essen, Germany, December 12, 2007.

Ray R. Larson. "Logistic Regression and EVI's for Book Search and the Heterogeneous Track." Paper and Presentation for the INEX (XML Retrieval) Workshop, Dagstuhl, Germany, December 18, 2007.

Ray R. Larson. "Bringing Lives to Light: Biography in Context" Presentation at WebWise, Miami, March 7, 2008

Michael Buckland. "Hybrid Models Should be the Norm. Why?-- And Why is it Unlikely?" California Association of Museums Annual Conference, Fresno Feb 26-27, 2008. Session on Museums and Libraries: A catalog of knowledge, An inventory of ideas. Feb 27, 2008.

Michael Buckland. "Documents, Documentation, Agendas." Keynote address, Document Academy Conference, Madison, WI, Mar 28, 2008.

Ray R. Larson, Fredric Gey and Ryan Shaw. "Biography in Context." Survey Research Center, 2538 Channing Way, Berkeley, CA. May 31, 2008

Michael Buckland. "Making contextual resources accessible for digital resources." *JISC-CNI Meeting: Transforming the User Experience, Belfast, Northern Ireland, July 10-11, 2008*. July 11, 2008.

Michael Buckland & Ryan Shaw. "4W vocabulary mapping across diverse reference genres." *International Society for Knowledge Organisation, Annual Conference*, Montreal, August 5-8, 2008.

Michael Buckland. "Library reference service in a digital environment." *Library and Information Science Research* 30, no 2 (2008): 81-85. http://people.ischool.berkeley.edu/~buckland/libref.pdf

Patricia Frontiera, Ray R. Larson and John Radke. "A comparison of geometric approaches to assessing spatial similarity for GIR". *International Journal of Geographical Information Science*, v. 22, no. 3, pp.337 – 360, 2008.

Ryan Shaw and Ray R. Larson. "Bringing Lives to Light: Lives and Event Representation in Temporal and Geographic Context". *Proceedings of the Joint ACM/IEEE Conference on Digital Libraries*, Pittsburgh, PA June 16-20, 2008.

Ray R. Larson. "Geographic IR and Visualization in Time and Space." *ACM SIGIR 2008: Thirty-first annual international ACM SIGIR Conference on Research and Development in Information Retrieval*. 20-24 July 2008, Singapore. P. 886, 2008

Ryan Shaw and Ray R. Larson. "Event Representation in Temporal and Geographic Context" In B. Christensen-Dalsgaard et al. (Eds.): *European Conference on Digital Libraries (ECDL 2008)*, LNCS 5173, pp. 415–418, 2008.

Michael Buckland, Ray R. Larson, and Ryan Shaw. "Finding and Providing Context Online", *Proceedings of the American Society for Information Science and Technology (ASIS&T) Annual Meeting* 2008.

Michael Buckland. "Laptops and the '9 to 5' problem: Providing reference library service in an online environment." University of Tromsø, Sept 24, 2008; also Tromsø Telemedicine Lab, Sept 25, 2008.

Michael Buckland. "The educational role of the library in a digital environment. Part I: Library reference resources when digital. Part II: Library design for learning." *Forum for Advanced Studies in Arts, Languages and Theology (SALT)*, Uppsala University, Sweden, Oct 2, 2008.

Fredric Gey, Ryan Shaw, Ray Larson, Michael Buckland, Barry Pateman and Dan Melia. "Marking Up Cultural Materials for Time and Geography". *Workshop on Information Access to Cultural Heritage*, Aarhus, Denmark, September 18, 2008.

Fredric Gey. "Multigenre Search Using Common Geography", presentation at the *HEIR (Hildesheim Evaluation and Information Retrieval) Workshop*, September 25, 2008, University of Hildesheim, Germany.

Ryan Shaw. "Event gazetteers for navigating humanities resources." In Proceedings. of the *2nd PhD Workshop on Information and Knowledge Management* (Napa Valley, CA) Oct. 2008.

Ryan Shaw and Michael Buckland. "Open identification and linking of the four Ws". In *Proc. of the 9th International Conference on Dublin Core and Metadata Applications* (Berlin, Germany) Oct. 2008.

Fredric Gey, "History as Events in Time and Space: Biography in Context" October 14, 2008. Presentation at the University of Hildesheim, Germany.

Fredric Gey. "Vocabulary Statistics Time and Geography" (Presentation at GESIS Leibniz Institute for the Social Sciences, Bonn, Germany) October 24, 2008.

Ray R. Larson. "Cheshire at GeoCLEF 2008: Text and Fusion Approaches for GIR". CLEF 2008 Notebook Paper. (Oct. 2008)

Fredric Gey, R Shaw, R Larson, and B Pateman, "Biography as Events in Time and Space", *Proceedings of ACM Geographic Information Systems (GIS) 2008 Conference*, (Irvine, California), Nov 5-7, 2008.

Michael Buckland, Ray R. Larson, and Ryan Shaw. "Finding and Providing Context Online", *Proceedings of the American Society for Information Science and Technology (ASIS&T) Annual Meeting,* Nov. 2008.

Ray R. Larson. "Ad hoc and Book XML Retrieval with Cheshire." INEX Notebook paper, Dagstuhl, Germany (Dec. 2008).

Michael Buckland. "Laptops and Libraries: Decentralized Access to Explanatory Resources". *International Joint GIS-IDEAS and PNC/ECAI Conference*, Hanoi, Vietnam, Dec 6, 2008.

Michael Buckland. "Access to Reference Resources in a Digital Environment". University of South Florida, Feb 3, 2009.

Michael Buckland. "The Electronic Cultural Atlas Initiative". University Libraries, Berkeley, Feb 5, 2009.

Jeanette Zerneke, "From Historical GIS to Seeing History: The Evolution of Digital Humanities", Visualizing the Past Workshop, University of Richmond: February 20-21, 2009.

Michael Buckland. "Empowering Readers to Find Explanations: A What - Where - When - Who Approach". *Electronic Cultural Atlas Initiative - Computer Applications in Archaeology Joint Conference, Williamsburg, "Making History Interactive," March 22-27, 2009*.

Matthew Holmberg. "Context and Relationships: Developing Electronic Research Tools for Irish Studies." *Joint meeting of the California Celtic Colloquium and the Celtic Studies Association of North America*, Berkeley, March 15, 2009.

Ray R. Larson and Ryan Shaw. "Mapping Life Events: Temporal and Geographic Context for Biographical Information". Proceedings of the ACM/IEEE Joint Conference on Digital Libraries 2009 (JCDL 2009), P. 471.

Ray R. Larson." Adhoc and Book XML Retrieval with Cheshire". IN: S. Geva, J. Kamps, and A. Trotman (Eds.): *INEX 2008*, Springer LNCS 5631, pp. 152–163, 2009.

Ray R. Larson. "Logistic Regression for Metadata: Cheshire Takes on Adhoc-TEL". In: C. Peters et al. (Eds.): *CLEF 2008*, Springer LNCS 5706, pp. 38–41, 2009.

Ray R. Larson. "Back to Basics – Again – for Domain – Specific Retrieval ". In: C. Peters et al. (Eds.): CLEF 2008, Springer LNCS 5706, pp. 203-206, 2009.

Thomas Mandl, Paula Carvalho, Giorgio Maria Di Nunzio, Fredric Gey, Ray R. Larson, Diana Santos, and Christa Womser-Hacker. "GeoCLEF 2008: The CLEF 2008 Cross-Language Geographic Information Retrieval Track Overview". In: C. Peters et al. (Eds.): *CLEF 2008*, Springer LNCS 5706, pp 808-821, 2009.

Ray R. Larson. "Cheshire at GeoCLEF 2008: Text and Fusion Approaches for GIR". In: C. Peters et al. (Eds.): *CLEF 2008*, Springer LNCS 5706, pp. 830-837, 2009.

Ray R. Larson. "Geographic Information Retrieval and Digital Libraries". In: M. Agosti et al. (Eds.): *ECDL 2009*, Springer LNCS 5714, pp. 461–464, 2009.

Ryan Shaw, Raphael Troncy, and Lynda Hardman. *LODE: Linking Open Descriptions of Events*. (August 3, 2009). School of Information. Paper 2009-036.

Michael Buckland, Ray Larson and Ryan Shaw. "Finding Explanations Online: Easy Queries AND Trustworthy Answers". Proceedings of the American Society for Information Science and TechnologyAnnual Meeting, Vancouver, BC, 2009.

Ray R. Larson. "Bringing Lives to Light: Biography in Context". Presented at the Berkeley Ischool, Kyoto University Workshop. September 2009.

Ray R. Larson. "Temporal and Geographic Context for Digital Books." Presentation at the ECDL Books Online Workshop. September 2009.

Michele R. Ramos. *Biography Light Ontology: An Open Vocabulary For Encoding Biographic Texts*. Technical Report. November 2009.

Ryan Shaw. "From Facts to Judgments: Theorizing History for Information Science". *Bulletin of the American Society for Information Science and Technology*, December/January 2010. http://www.asis.org/Bulletin/Dec-09/Bulletin_DecJan10_Final.pdf

Michael Buckland and Michele Renee Ramos. "Events as a Structuring Device in Biographical Mark-up and Metadata". *Bulletin of the American Society for Information Science and Technology*, December/January 2010. http://www.asis.org/Bulletin/Dec-09/Bulletin_DecJan10_Final.pdf

Conclusion

The Bringing Lives to Light project has been a very productive and successful research project, and in turn has opened up a range of new research possibilities. The PI and all of the collaborators on the projects wish to thank IMLS for their support of this interesting and fruitful work.

Respectfully Submitted,

Ray R. Larson
Principal Investigator