

Gerry McKiernan

Open archives initiative service providers. Part II: social sciences and humanities

This is the second E-profile that describes Open Archives Initiative service providers. The first in the series reviewed select services that offer access to a range of science and technology repositories, while the third will focus on those that provide access to a broad array of subjects and disciplines.

AmericanSouth.org

AmericanSouth.org seeks to “make material crucial to the understanding of the Southern experience accessible to all citizens through the creation of a collaborative digital collection of Southern history and culture.” It is expected that AmericanSouth.org will “significantly improve cultural literacy and enhance library services and scholarly communication by expanding opportunities for sharing information and removing geographic and other barriers to collections” (www.solinet.net/preservation/preservation_tmpl.cfm?doc_id=645). The AmericanSouth.org project includes the following partners:

- Auburn University;
- Emory University;
- Louisiana State University;
- University of Florida;
- University of Georgia;
- University of Kentucky and the Kentucky Virtual Library;
- University of North Carolina at Chapel Hill;
- University of Tennessee at Knoxville; and
- Vanderbilt University.

Currently, the AmericanSouth.org project consists of five components with the following scope and coverage:

- “Research Collections” (provides “access to the archival collections of more than three dozen universities, colleges, and museums”);
- “Articles” (contains “peer-reviewed articles written by experts on a wide range of Southern Studies topics”);
- “Weblinks” (is a browsable collection of research-quality Web resources about the American South);
- “Discussions” (provides access to electronic discussion forums relating to the American South and its culture); and
- “Encyclopedia of Southern Culture” (digital full-text access to the standard encyclopedia about the South) (forthcoming).

As of mid-October 2003, nearly 30,000 (29,559) source items were available from the AmericanSouth.org service in its “Research Collection”. Among the repositories and resources accessible from the “Research Collections” are:

- American Routes (Public Radio International) (www.amroutes.com).
- Digital Library of Georgia (dlg.galileo.usg.edu) (University System of Georgia).
- Documenting the American South (University of North Carolina at Chapel Hill) (docsouth.unc.edu).
- Electronic Text Center (University of Virginia Library) (etext.lib.virginia.edu).
- Florida Environments Online (State University System of Florida) (palmm.fcla.edu/feol/).
- Florida Heritage Collection (State University System of Florida) (palmm.fcla.edu/fh/).
- John G. Towers Papers (Southwestern University) (www.southwestern.edu/library/tower/).

- Southern Changes Online (Emory University) (chaucer.library.emory.edu/htdocs/schanges.html).
- Tennessee Documentary History (University of Tennessee) (diglib.lib.utk.edu/dlc/tdh/).
- Valley of the Shadow (Virginia Center for Digital History, University of Virginia) (valley.vcdh.virginia.edu).

Currently, the “Research Collections” collection service can be browsed by archive, or searched separately or concurrently with other AmericanSouth.org components. When browsed, entries will be displayed in a brief format that provides a full or partial title of the source (e.g. “Charles B. Carter to Eliza Carter, April 26, 1862”, the archive name (e.g. “VCDH Valley of the Shadows”) and an OAI identifier (e.g. “oai:valley.vcdh.virginia.edu:A0100”). In addition, most brief records also include an unidentified field that provides data or information about the resource (for example, a truncated statement about the nature of the resource (e.g. “Carter writes to his wife about life in the Camp Chase Prison . . .”). The associated full record is displayed by clicking the title entry.

A typical full record in AmericanSouth.org includes most applicable Dublin Core Metadata Set elements (dublincore.org/documents/dces/) in the following sequence: “Type”, “Subject”, “Language”, “Publisher”, “Source”, “Contributor”, “Author”, “Rights”, “Identifier”, “Description” [Abstract], “Title”, and “Date”. In addition, other fields specific to the AmericanSouth.org full record are interspersed, notably “Archive” (e.g. “VCDH Valley of the Shadows”), “ID” (e.g. “oai:valley.vcdh.virginia.edu:A0100”), “Discovery” (e.g. “1862-04-26”) and “Sets”. Generally the identifier is hotlinked and provides direct access to an item.

AmericanSouth.org currently offers one search option: a free text, basic search that is intended to allow the user to search all or one of its site resource groups by accepting the default (“Search All”) or by selecting from available options (i.e. “Search Collections”, “Search Articles”, “Search Web Links”, or “Search FAQ”) from a pull-down menu. Although the user can specify a particular resource group, all groups are searched concurrently at the present time. Although not stated, users can perform phrase searching by enclosing terms within double quotes (e.g. “Slave Narratives”), truncate terms using an asterisk (“*”) (e.g. “potato*”), and nest terms in a Boolean statement (e.g. “(Irish or Ireland) and potato*”).

Upon execution, matching results are displayed by resource group type (i.e. “Special Collections”, “Articles”, “Web Links”, and “Frequently Asked Questions”) in individual scrollable windows. Above the grouped search results is a summary listing the overall results (for example, “16 SPECIAL COLLECTION RECORDS FOUND”, “2 ARTICLES FOUND”, “0 WEB LINKS FOUND”, “0 FAQ ENTRIES FOUND”). In this context, the “SPECIAL COLLECTION” is the AmericanSouth.org “Research Collections”. Results from the “SPECIAL COLLECTION” search will be displayed in the same brief format as entries displayed in browsing the “Research Collections”. A function that will allow users to select and deposit items into a “My Collection” set is currently under development.

Several additional collections are planned for harvesting for the AmericanSouth.org service and include:

- Sam Fleming Southern Civilization Collection (Vanderbilt University Libraries) (www.library.vanderbilt.edu/speccol/collections.shtml).
- Southern Civilization Collection (Vanderbilt University Libraries) (www.library.vanderbilt.edu/speccol/collections.shtml).
- Southern Folklife Collection (University of North Carolina at Chapel Hill) (www.lib.unc.edu/mss/sfc1/).
- Transforming America: Alabama’s Role in the Civil Rights Movement (Auburn University) (diglib.auburn.edu/transam/)

AmericanSouth.org is one of seven demonstration projects funded in 2001 by The Andrew W. Mellon Foundation to test the application of the Open Archives Initiative Metadata Harvesting Protocol (OAI-MHP) (Waters, 2001). Project management responsibilities for AmericanSouth.org were subsequently transferred from the original project group, Southeastern Library Network, Inc. (SOLINET) (www.solinet.net), to Emory University. At Emory, it was merged with another Mellon-funded project, MetaArchive.org (www.metaarchive.org) to form an umbrella project known as the MetaScholar Initiative (www.metascholar.org) (Halbert, 2003). Martin Halbert (mhalber@emory.edu), Director for Library Systems, Emory University, serves as the principal investigator for AmericanSouth.org and MetaArchive.org, both cooperative projects of Emory University and the Association of Southeastern Research Libraries (ASERL).

From within MetaArchive.org, users can search select AmericanSouth.org collections, as well as the MetaArchive.org collections of archival papers of major political leaders and theological records, using a basic or “Extended Search” interface. Currently, MetaArchive.org provides access to The Church in the Southern Black Community, 1780-1925 (Documenting the American South), the Pitts Theology Library collections (Emory University) (www.pitts.emory.edu), the Senator Sam Nunn papers (Emory University) (sage.library.emory.edu/collection-0800.html), the Senator John G. Towers papers, and collections of the General Commission on Archives and History for The United Methodist Church (www.gcch.org).

Open Language Archives Community (OLAC)

The Open Language Archives Community (OLAC) (www.language-archives.org) “is an international partnership of institutions and individuals who are creating a worldwide virtual library of language resources by: developing consensus on best current practice for the digital archiving of language resources, and developing a network of interoperating repositories and services for housing and accessing such resources” (see also

Simons and Bird, 2003a). In OLAC, a language resource is defined as:

... any kind of DATA, TOOL or ADVICE pertaining to the documentation, description or analysis of a human language ... [O]LAC metadata can be used to describe any kind of language resource. Language resources may be digital or non-digital, published or restricted. A language archive is any collection of language resources and their resource descriptions (www.language-archives.org/documents/faq.html#2).

From the OLAC perspective, “... individuals who use and create language resources are looking for three things: data, tools, and advice.”

By DATA we mean any information that documents or describes a language, such as a published monograph, a computer data file, or even a shoebox full of handwritten index cards. The information could range in content from unanalyzed sound recordings to fully transcribed and annotated texts to a complete descriptive grammar. By TOOLS we mean computational resources that facilitate creating, viewing, querying, or otherwise using language data. Tools include not just software programs, but also the digital resources that the programs depend on, such as fonts, stylesheets, and document type definitions. By ADVICE we mean any information about what data sources are reliable, what tools are appropriate in a given situation, what practices to follow when creating new data, and so forth (Bird and Simons, 2003).

As of late 2002, approximately 30,000 records had been harvested from 24 repositories (Simons and Bird, 2003a). As of mid-October 2003, there were twenty seven repositories whose metadata was harvested by the service (www.language-archives.org/archives.php4).

Among the common and uncommon harvested repositories (www.language-archives.org/organization.html) are:

- Academia Sinica Balanced Corpus of Modern Chinese (www.sinica.edu.tw/SinicaCorpus/).
- Archive of the Indigenous Languages of Latin America (www.ailla.org).
- Archives of the Alaska Native Language Center (archives.anlc.uaf.edu).

- ASEDA: the Aboriginal Studies Electronic Data Archive (coombs.anu.edu.au/SpecialProj/ASEDA/).
- Ethnologue: Languages of the World (www.ethnologue.com).
- Natural Language Software Registry (registry.dfki.de).
- Perseus Digital Library (www.perseus.tufts.edu/).
- Rosetta Project (www.rosetta-project.org:8080/live/).
- TalkBank (talkbank.org).
- Tibetan and Himalayan Digital Library (iris.lib.virginia.edu/tibet/).

The OLAC Metadata Set (OLACMS) (www.language-archives.org/OLAC/metadata.html) is “the set of metadata elements that members of OLAC have agreed to use for describing language resources.” The OLACMS consists of all 15 Dublin Core Metadata Set elements (dublincore.org/documents/dces/) as well as refinements and qualifications designed for describing fundamental properties of language resources. Specifically, two elements have been added: a “Subject.language” element and a “Type.linguistic” element. “Whereas the Language element in the (Dublin Core) standard identifies the language in which the resource is written or spoken, a Subject.language element is added to identify the language which a resource is about. Whereas the Type element (in the Dublin Core) identifies what type of thing the resource is generically (such as sound or image or readable text), a Type.linguistic element is added to identify what type of thing the resource is from a linguistic point of view (such as a text or a lexicon or a grammatical description)” (Simons and Bird (2003b) – see also Bird and Simons (2003)).

The OLAC service currently offers two primary search options: a basic search and an ‘Advanced Search.’ From the basic search interface page (linguistlist.org/olac/), the user can perform a keyword search in one of four fields: title, description, creator and contributor. After execution, the results are displayed in a brief record format that provides the title of a matching resource (e.g. “GAELIC, IRISH, a language of Ireland”) and the data provider (‘archive’) name (e.g. “Ethnologue: Languages of the

World”). On clicking the hotlinked title entry, a full record is retrieved, displaying more specific data and information about the resource, an example, is shown below:

- (1) “General Description”;
 - “Title” (e.g. “GAELIC, IRISH: a language of Ireland”).
 - “Archive” (e.g. “Ethnologue: Languages of the World”).
 - “Description” (e.g. “A page from the Web edition of Ethnologue: Languages of the World (14th edition) giving basic facts about the language and where it is spoken”).
 - “Date” (e.g. “2000”).
 - “Format” (e.g. “text/html”).
- (2) “Linguistic Description”; and “Subject” (e.g. “(x-sil-GLI)”);
- (3) “Access Description”:
 - “Publisher” (e.g. “SIL International (www.sil.org)”).
 - “Identifier” (e.g. “http://www.ethnologue.com/show_language.asp?code=GLI”).

Identifiers are hotlinked to their source item or to a record, or entry in a source archive.

In an “Advanced Search”, a user can perform a keyword search in a specific incorporated collection (“Archive”) by selecting those available from a pull-down menu. In addition, the user can perform a “Title” search or search by “Creator/Contributor”. A search can also be limited by “Subject Language” or linguistic “Type” (see above) (see Figure 1). In the “Subject Language” the more common modern languages (e.g. “French”) or ancient languages (e.g. “Greek, Ancient”) can be selected from a scrollable list; a query can be limited by languages other than those listed by selecting and searching the “More Languages” button found adjacent to the menu. In addition, a search can be limited by a “Type” element. While not specifically defined in OLAC, a “Type” can be selected from a separate scrollable menu that includes the items shown in Table I).

The Open Language Archives Community invites all members of the language resources community to

contribute to the OLAC initiative. To ensure that participation is as broad as possible, OLAC offers three approaches to becoming an OLAC data provider (www.language-archives.org/documents/implement.html) (see Table II and also Simons and Bird (2003a)).

Through its OLAC Repository Editor (linguistlist.org/olac/ore/index.html), OLAC enables every researcher and archive to become a data provider by submitting information and data about available language-oriented resources (e.g. field notes, grammatical sketches, lexicons, unpublished papers, etc.) using simple Web-based forms.

The services on the OLAC site have been developed by Steven Bird (sb@cs.mu.oz.au), Associate Professor, Computer Science and Software Engineering, University of Melbourne, Australia, and Senior Research Associate at the Linguistic Data Consortium, University of Pennsylvania; Gary F. Simons (gary_simons@sil.org), Associate Vice President for Academic Affairs at SIL International, Dallas, Texas; and Haejoong Lee (haejoong@ldc.upenn.edu), the Linguistic Data Consortium.

The OLAC service is hosted on the LINGUIST List Web site (linguistlist.org) – “the world’s largest online linguistic resource”. The aim of the list is “to provide a forum where academic linguists can discuss linguistic issues and exchange linguistic information. LINGUIST is operated by a staff of linguistics professors and graduate students affiliated with Eastern Michigan University, Wayne State University, and [the] University of Arizona.” The LINGUIST Web site also provides access to a variety of Linguistics information resources and sources. Individuals specifically interested in language resource development and archiving can subscribe to the OLAC-General electronic discussion list (lists.linguistlist.org/archives/olac-general.html) from the LINGUIST Web site.

Support for the core infrastructure of the Open Language Archives Community service was provided by the National Science Foundation through grants made to the University of Pennsylvania, Carnegie Mellon University, and Wayne State University. The harvester software developed for the OLAC project is open source and

Figure 1
 Screen print of Open Language Archives Community (OLAC) “Advanced Search” interface

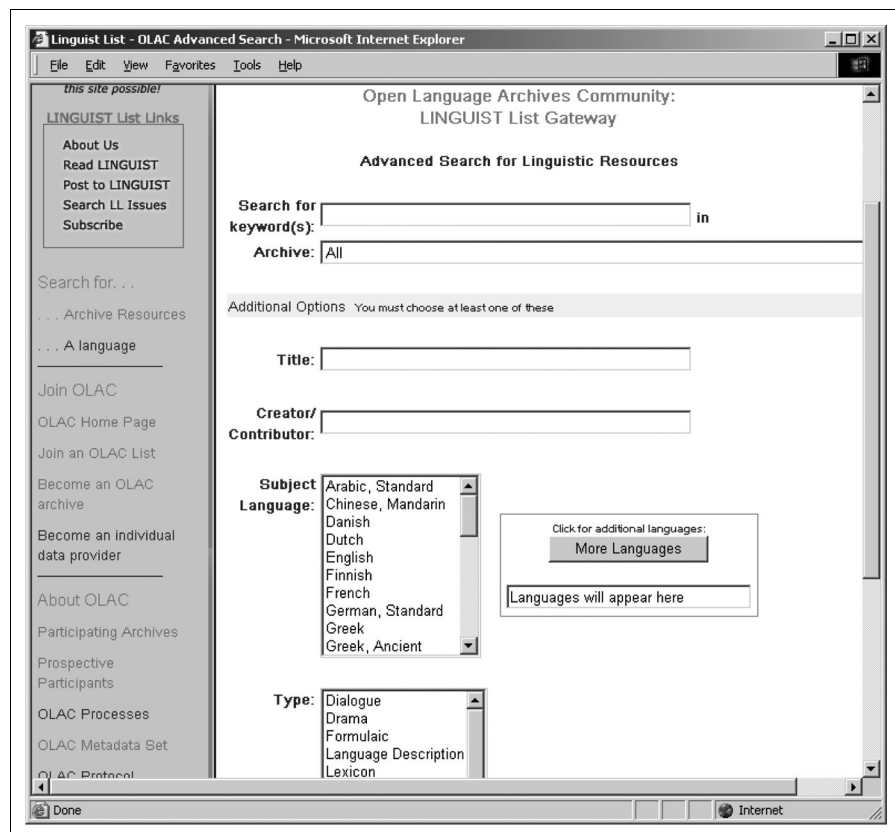


Table I

Dialogue	Drama	Formulaic
Language description	Lexicon	Ludic
Narrative	Oratory	Primary text
Report	Singing	Unintelligible speech

available from SourceForge.net (source forge.net).

Sheet Music Consortium

The Sheet Music Consortium is a group of libraries that seeks to build “an open collection of digitized sheet music” for use by musicians as well as non-musicians, employing the OAI-PMH (Davison, 2003; see also Dunn and Riley, 2003). Currently, metadata for the Sheet Music Consortium service (digital.library.ucla.edu/sheetmusic/) (see Figure 2) is harvested from the following collections:

- *Digital Archive of Popular American Music (University of California,*

Los Angeles) (digital.library.ucla.edu/apam/):

The Archive of Popular American Music is a non-circulating research collection covering the history of popular music in the United States from 1790 to the present. The collection is one of the largest in the country, numbering almost 450,000 pieces of sheet music, anthologies, and arrangements for band and orchestra, and 62,500 recordings on disc, tape, and cylinder. The Digital Archive of Popular American Music is an initiative designed to provide access to digital versions of the sheet music, and performances of the songs now in the public domain.

- *Historic American Sheet Music Project (Duke University) (scriptorium.lib.duke.edu/sheetmusic/):*

The Historic American Sheet Music Project provides access to digital images of 3,042 pieces from the collection, published in the United States between 1850 and 1920 (see also Storey, 2003).

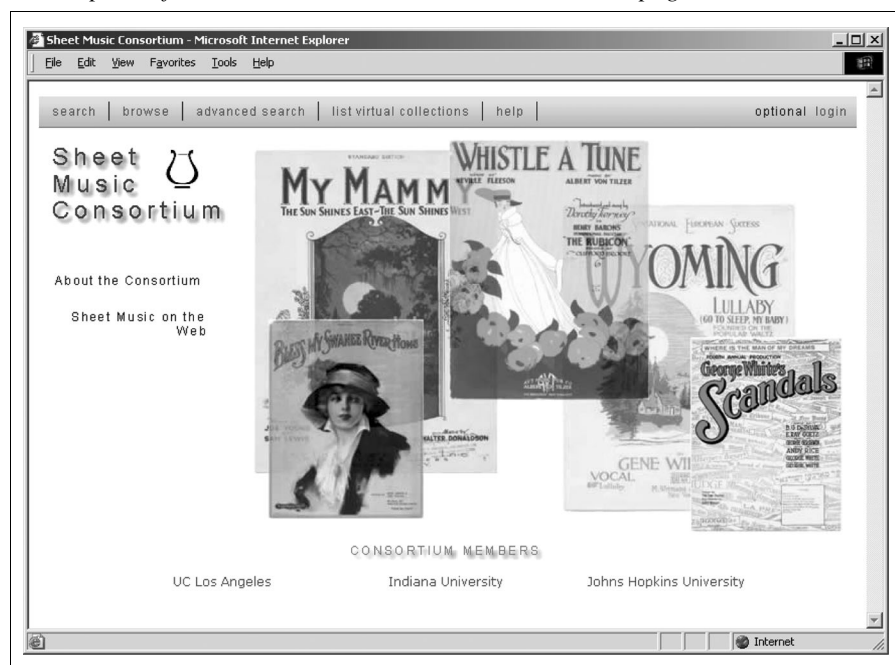
- *Lester S. Levy Collection of Sheet Music (Johns Hopkins University) (levysheetmusic.mse.jhu.edu):* The Lester S. Levy Collection of Sheet Music “contains over 29,000 pieces of music and focuses on popular American music spanning the period 1780 to 1960.
- *Music for the Nation: American Sheet Music (Library of Congress) (memory.loc.gov/ammem/mussmhtml/mussmhome.html):* Music for the Nation: American Sheet Music contains more than 62,500 pieces of historical sheet music registered for copyright: more than 15,000 registered during the years 1820-1860 and more than 47,000 registered during the years 1870-1885. Included are popular songs, operatic arias, piano music, sacred and secular choral music, solo instrumental music, method books and instructional materials, and music for band and orchestra.
- *Sam DeVincent Collection of American Sheet Music (Indiana University) (www.indiana.edu/~liblilly/devincent.shtml):* The Sam DeVincent Collection of American Sheet Music “contains approximately 24,000 pieces of sheet music, songbooks, and folios.

As of late July 2003, nearly 100,000 (96,406) source items were available from the Sheet Music Consortium service, with half (47,528 records) harvested from the Music for the Nation: American Sheet Music collection of the Library of Congress, one-fifth (17,698) from the Historic American Sheet Music Project collection at Duke University, another fifth (17,417) from the Sam DeVincent Collection of American Sheet Music, and the remainder from the other two source collections, The Lester S. Levy Collection of Sheet Music (11,590) and the Digital Archive of Popular American Music (2,173) (digital.library.ucla.edu/sheetmusic/OAIProject.html).

The Sheet Music Consortium service currently offers two search options: “Search” and “Advanced Search.” From the “Search” interface page, the user can perform a keyword search in any of 14 unqualified Dublin Core (dublincore.org) fields, where available

Table II*Description of three major approaches for participating in the OLAC service*

Approach	Description	Circumstances
Conventional repository	The data provider implements a software interface to an existing catalog database that permits metadata to be harvested via an HTTP-based protocol	Collection is cataloged in a database Web server available Programmer who can implement a metadata harvester available
Static repository	The Static Repository Gateway is an OLAC service that provides the harvesting protocol for metadata repositories that are expressed as XML documents	No catalog database Provider can use an XML editor to create one, or existing catalog database and the provider can export it to XML
OLAC Repository Editor (ORE)	The OLAC Repository Editor is a form-based editor on the OLAC Web site that allows a user to construct a small metadata repository that will be served by the OLAC Static Repository gateway	No catalog database Provider has no local Web site or XML capacity

Figure 2*Screen print of the Sheet Music Consortium service homepage*

(see below) as well as by collection name (see above). If desired, search results can be limited to records that include a link to digitized sheets (“Digitized Sheet Music Only”). Users can perform phrase searches by

enclosing terms in double quotes, use common Boolean operators (AND, OR, NOT), and use a special character (“%”) as a wild card or for truncation. A query search will be expanded to include all terms with similar spelling

when a question mark symbol (“?”) is placed before a term or a portion of a term. By default, search results will be displayed twenty records to a page, although the user can decrease or increase the number (“10”, “40”, “100”, “200”) by selecting from a drop-down menu.

After execution, search results are displayed in a brief format that includes the following:

- “Title” (e.g. “More Love Than Your Love”);
- “Creator” (e.g. “Music by Arthur Schwartz. Lyrics by Dorothy Fields. Book by Herbert and Dorothy Fields”);
- “Publisher” (e.g. “New York: Edwin H. Morris & Company, Inc., 1619 Broadway 1954”); and
- “Collection” (e.g. “Johns Hopkins University Levy Collection”).

Records can include more than one creator field, reflecting different creative roles (e.g. arranger, composer, lyricist, performer). Beneath each brief record are links to three (or four) actions. When available, the first provides direct access to the fully digitized version of the source music sheets (“access online”), the next allows the user to display an item’s full record (“more info”), while the two remaining options enable the user to “add” the item to a “virtual collection” or annotate a record (“add with a note”).

A full record within the Sheet Music Consortium service includes many (or all) of the following fields (digital.library.ucla.edu/sheetmusic/project.jsp#standards), all but the last (“Collection”) a Dublin Core field (see Table III).

The last field, “Collection”, provides the name of the respective data provider (e.g. “Johns Hopkins University Levy Collection”). As in the brief record format, within the full record display the user can also annotate a record, select it, or access the source item.

As noted, clicking the “Add” button found beneath a brief record display will add that record to a ‘virtual collection’ during the course of a current session. A virtual collection “is an application that allows any user to save personal collections of sheet music

Table III*List of field labels in a Sheet Music Consortium service in a complete full record*

Title	Creator	Subject	Description
Publisher	Date	Type	Format
Identifier	Source	Language	Relation
Coverage	Rights	Collection	

and attach notes to the records. Added notes do not change the original record, nor do they become a part of it, except within the collection made by the user.” Although any user can create a virtual collection, only registered users can “lock” their collections, “protect” the notes associated with annotated records, and control access (digital.library.ucla.edu/sheetmusic/help.jsp#virtual_collections). In the registration process, an individual is requested to provide a user-chosen “login ID”, password, name, and e-mail address.

Items selected within a virtual collection will be retained only for the course of a session unless the collection is formally named and saved within the service site. Alternatively, a user can e-mail the contents of a virtual collection; the e-mail record is brief, providing the title, creator(s), and an associated URL (e.g. “<http://digital.library.ucla.edu/sheetmusic/librarian?ITEMID=59413>”) which links to the full record for the source item.

An “Advanced Search” allows the user to query the consortium database by one of four major fields (i.e. “Composer or Lyricist”, “Title”, “Subject”, or “Publisher”) (as well as “Keyword”) by selecting from the field option from a pull-down menu. Terms can be combined in a Boolean statement (AND, OR, NOT) with other terms from a secondary query by accepting or changing the options from a logical operator pull-down menu. A third and fourth query box are also available in the advanced search interface with identical Boolean options. A search can be limited by date range or consortial collection (e.g. “Johns Hopkins University Levy Collection”), sorted by title (default) or date, and/or restricted to records that include a link to the original source item. The Sheet Music Consortium service can also be browsed by the first letter of an item’s title, date (or date

range), and/or collection, and results sorted by title (default) or date.

In addition to providing access to several major sheet music collections through a common interface, the Sheet Music Consortium also includes a link to a directory of these and other individual collections (www.lib.duke.edu/music/sheetmusic/collections.html), as well as information sources about sheet music (digital.library.ucla.edu/sheetmusic/OAIProject.html).

The UCLA Digital Library Program (digital.library.ucla.edu) hosts the metadata harvested from participating collections in the Sheet Music Consortium (UCLA Digital Library, 2002). Brown University, which participated in usability studies of a prototype user interface for the service, is expected to formally join the consortium in the near future. The consortium welcomes participation by other sheet music collections; inquiries should be addressed to Stephen H. Schwartz (shs@library.ucla.edu), chair of the project team, and Head of Systems Development, UCLA Library. Currently, the project is focused on metadata and/or images of sheet music only; it is not presently interested in music scores or audio collections. Although it will accept metadata only, it seeks to build an international collection of digitized sheet music for online viewing.

UIUC Digital Gateway to Cultural Heritage Materials

The UIUC Digital Gateway to Cultural Heritage Materials (oai.grainger.uiuc.edu) was created to serve as “a portal to facilitate access to scholarly cultural heritage information” and as an environment to “investigate the efficacy of harvesting and aggregating metadata using the Open Archives Initiative (OAI) Protocol for Metadata Harvesting (PMH)” (see also Shreeves *et al.* (2003)).

In March 2003, the UIUC Digital Gateway to Cultural Heritage Materials service contained more than 1.1 million records with metadata describing more than 2.5 million discrete items held in approximately 500 museums, archives, historical societies, and academic and public libraries, located or aggregated in 39 repositories. Among the common and uncommon harvested repositories are (oai.grainger.uiuc.edu/About_Collections.htm):

- Alex Catalogue of Electronic Texts (www.infomotions.com/alex/);
- American Memory (Library of Congress) (lcweb2.loc.gov/amhome.html);
- Consortium of Museum Intelligence (CIMI) Demonstration Repository (cimi.adlibsoft.com/dct/data/dct.html);
- David Rumsey Map Collection (www.davidrumsey.com);
- Formations (formations2.ulst.ac.uk);
- Heritage Colorado (www.cdheritage.org);
- Indiana University Digital Library Program (www.dlib.indiana.edu);
- Open Video Project (www.open-video.org);
- Schoenberg Center for Electronic Text & Image (SCETI) (University of Pennsylvania (dewey.library.upenn.edu/sceti/); and
- University of Michigan Digital Library Text Collections (www.umdl.umich.edu).

The UIUC Digital Gateway to Cultural Heritage Materials service offers two major search options: “Simple Search” and “Advanced Search.” In a “Simple Search” the user can perform a keyword search in any of the incorporated collections (“anywhere in all collections”). A search can be limited to records that provide links to associated digitized resources (“online primary sources”).

After execution, search results are displayed in a “short” format that includes the following labeled fields: “Title”, “Author/Artist”, “Subject/Description”, “Online Access” (if available) and “Collection”. Beneath the last field (“Collection”) is a link to an item’s “full record” and a save function (“add to book bag”). By default, all retrieved records are

concatenated in a combined set (“All Types of Material”), grouped by record type. Current record types include “Images and Video”, “Museum and Archives Collection”, and “Text, Sheet Music, and Websites”. On average, two-dozen records are displayed per page. The user can navigate to other pages by selecting from a numbered sequence of pages listed above the results display, or proceed page-by-page by clicking the “Next page” option found to the far right of the sequence.

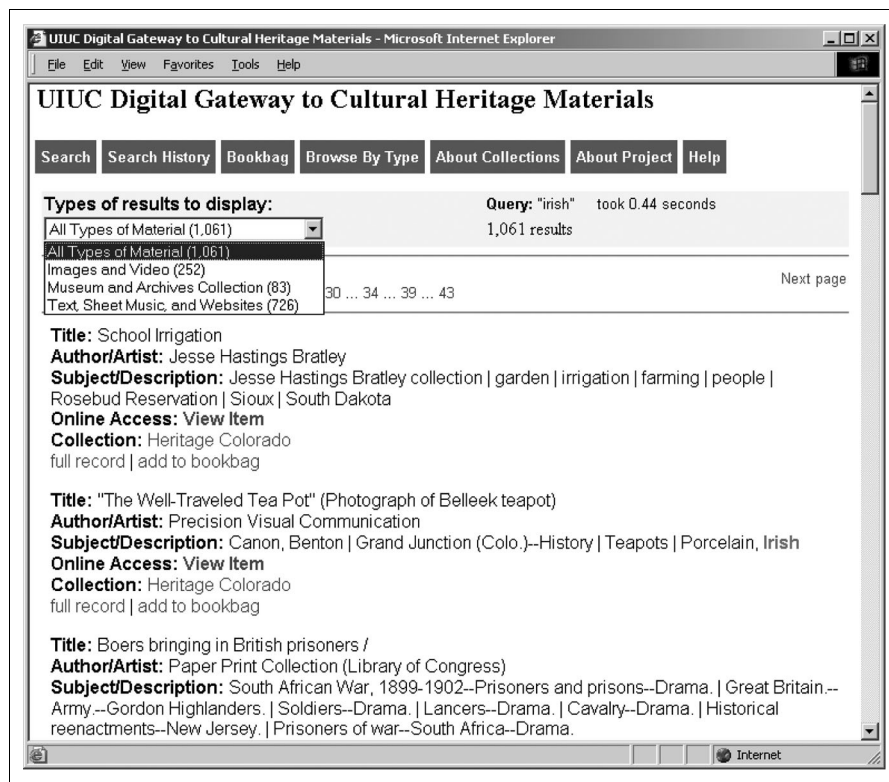
If desired, results can be displayed by a specific record group by choosing the record type option from a pull-down menu found in the upper left-hand corner of a search results page (“Types of results to display”). At this time sorting by a specific field, for example “Author/Artist”, is not currently offered. From within the short record, the user can access the source item directly by clicking the “View Item” option in the “Online Access” field, if available (see Figure 3).

The full record can be displayed by selecting the link associated with a particular record, and can include the following fields (oai.grainger.uiuc.edu/help.html#background):

- (1) Title (e.g. “The Well-Traveled Tea Pot” (Photograph of Belleek Teapot)).
- (2) Creator (e.g. “Precision Visual Communication”).
- (3) Subject, e.g.:
 - Canon, Benton;
 - Grand Junction (Colo.) – History;
 - Teapots;
 - Porcelain, Irish.
- (4) Publisher (e.g. “Museum of Western Colorado”);
- (5) Contributor, e.g.:
 - Donor: Carolyn Morton;
 - Researcher: David Sundal;
 - Cataloger: Wayne Randolph.
- (6) Description, e.g.
 - Master: Desk Scan 3500, Agfa Photo Look, tif, 355 ppi, RGB 24 bit, no compression, 29.6 MB.

Figure 3

Sample search results from the UIUC Digital Gateway to Cultural Heritage Materials service



- Thumbnail/enlargement: Microtek Scanmaker E6, jpeg, 72 ppi.
- Title supplied by researcher. 2 1/4" color transparency of white porcelain tea pot in pentagonal bamboo motif with gold painted detail. Length 21 cm; width 12cm; height 12.3cm. Wedding gift to Benton and Evelyn Canon, ca. 1865. First brought to Grand Junction, 1886. The tea pot was donated to the Museum in 1979).
- (7) Format/Quantity of Material (e.g. “web browser”);
- (8) Date, e.g.:
 - Master: 6/22/01;
 - Digital: 2/21/02;
 - 2001, 2 1/4 color transp).
- (9) Relation (e.g. “Tea pot donated with box”).
- (10) Type of Material (e.g. “image”).
- (11) Source (e.g. “Accession No. 1979.102.1”);
- (12) Online access/Unique ID (e.g. “http://www.gjhistory.org/dec/IMAGES2/Teaptpc.jpg”).
- (13) Rights (e.g. “Contact the Museum of Western Colorado for information concerning any copyright or other restrictions that may apply to the use or reproduction of this image”).

When online access is available for an item (“Online access/Unique ID”), a Uniform Resource Locator (URL) is hotlinked.

From within the same search page available for a “Simple Search”, the user also has the option of submitting and executing an “Advanced Search”. This option allows the user to search by “any field”, “author/artist”, or “title/subject” by selecting the field name from a pull-down menu. As with its sister service, the Open Archives Initiative Information in Engineering, Computer Science, and Physics (Grainger Engineering Library at the University of Illinois at Urbana-Champaign) (g118.grainger.uiuc.edu/engroai/), terms can be combined with others by entering these additional

terms into a secondary search box and selecting an appropriate field from an identical pull-down menu. A tertiary search box is also available. The user can combine the terms (or phases) using one of three Boolean operations: “also must contain” (AND), “or may contain” (OR), or “but not contain” (AND NOT).

The UIUC Digital Gateway not only allows the user to limit records that provide links to associated digitized resource as in the “Simple Search”, but also to a specific record type or range of dates. All record types will be retrieved by accepting the default for this option (“Types”: all checked boxes); a specific type can be selected by deselecting those types that are not of interest (“Images and Video”, “Museums and Archives Collection”, “Text, Sheet Music, and Web sites”).

A scroll menu allows users to search all dates (default) or limit to one or more specific periods by pointing, holding, and clicking those periods of interest. For the earlier periods, only a century range (e.g. “18th Century (1700-1799)”) is available; for more recent periods, more detailed periods are available (for example, the “20th Century (1900-1999)”) is further subdivided by the following periods: “1900-1924”, “1925-1949”, “1950-1974”, “1975-1999”).

As with its sister service, the UIUC Digital Gateway to Cultural Heritage Materials service allows the user to select and collect items by clicking the “add to bookbag” function found beneath each brief record. Within a “bookbag”, records are stored in brief record format, but can be displayed in full by clicking the “Long Display” button, or saved (currently in XML in the brief format only), by selecting the “Download Contents” button, within the collection.

The UIUC Digital Gateway to Cultural Heritage Materials project was funded by the Andrew W. Mellon Foundation (Cole *et al.*, 2003). Timothy W. Cole (t-cole3@ux1.cso.uiuc.edu), Mathematics Librarian, University of

Illinois at Urbana-Champaign, served as the lead principal investigator. Software developed by the project has been made available under an Open Source Initiative license (www.opensource.org).

ACKNOWLEDGEMENTS

The author is most grateful to the following individuals for granting permission to use screen images from their respective projects: Gary F. Simons, SIL International, Dallas, Texas; Stephen H. Schwartz, University of California, Los Angeles; and Timothy W. Cole, University of Illinois at Urbana-Champaign.

REFERENCES

Bird, S. and Simons, G. (2003), “Extending Dublin Core Metadata to support the description and language resources”, *Computing and the Humanities*, Vol. 37 No. 4, November, pp. 375-88. Also available at: http://arxiv.org/PS_cache/cs/pdf/0308/0308022.pdf (accessed 17 October).

Cole, T.W., Habing, T.G., Mischo, W.H., Prom, C., Sandore, B., Kaczmarek, J. and Shreeves, S. (2003), “Implementation of a scholarly information portal using the Open Archives Initiative Protocol for metadata harvesting”, University of Illinois at Urbana-Champaign, Final Report to the Andrew W. Mellon Foundation, 25 July, available at: http://oai.grainger.uiuc.edu/FinalReport/Mellon_Final_Report.doc (accessed 15 October).

Davison, S. (2003), “The Open Archives Initiative (OAI) Sheet Music Project: A Gateway to Sheet Music Collections on the Web”, PowerPoint presentation given at the 72nd Annual Meeting of the Music Library Association, Austin, TX, February 13, available at: <http://unitproj.library.ucla.edu/music/oaisheetmusic/mla.ppt> (accessed 12 October).

Dunn, J. and Riley, J. (2003), “The Open Archives Initiative and the Sheet Music Consortium”, PowerPoint presentation given at the Fall 2003 Brown Bag Workshop, Digital Library Program, School of Library and Information Science, Indiana University, Bloomington,

IN, October 10, available at: www.dlib.indiana.edu/workshops/oai/OAI-Oct2003.ppt (accessed 12 October).

Halbert, M. (2003), “The Metascholar Initiative: AmericanSouth.org and MetaArchive.org”, *Library Hi Tech*, Vol. 21 No. 2, pp. 182-98, available at: <http://callimachus.library.emory.edu/metascholar/Documents/MetaScholar.pdf> (accessed 17 October).

Shreeves, S.L., Kaczmarek, J.S. and Cole, T.W. (2003), “Harvesting cultural heritage metadata using the OAI protocol”, *Library Hi Tech*, Vol. 21 No. 2, pp. 159-69, available at: http://chaos.vtls.com/oai_docs/OAI_shreeves.pdf (accessed 14 October).

Simons, G. and Bird, S. (2003a), “Building an Open Language Archives Community on the OAI foundation”, *Library Hi Tech*, Vol. 21 No. 2, pp. 210-18, available at: <http://arxiv.org/ftp/cs/papers/0302/0302021.pdf> (accessed 17 October).

Simons, G. and Bird, S. (2003b), “The Open Language Archives Community: an infrastructure for distributed archiving of language resources”, *Literary and Linguistic Computing*, Vol. 18 No. 2, July, pp. 117-28, available at: <http://arxiv.org/ftp/cs/papers/0306/0306040.pdf> (accessed 16 October).

Storey, T. (2003), “University repositories: an extension of the library cooperative”, *OCLC Newsletter*, No. 261, July, pp. 8-9. Also available at: www.oclc.org/news/publications/newsletters/oclc/2003/261/n261.pdf (accessed 12 October 2003).

UCLA (2002), “UCLA Digital Library”, *UCLA Library News for the Faculty Online*, Vol. 17 No. 3, Fall, available at: www.library.ucla.edu/publications/facnews/vol17no3/apam.html (accessed 12 October).

Waters, D.J. (2001), “The metadata harvesting initiative of the Mellon Foundation”, *ARL Monthly Report*, No. 217, August, pp. 10-11, available at: www.arl.org/newsltr/217/waters.html (accessed 17 October).

Gerry McKiernan (gerrymck@iastate.edu) is a Science and Technology Librarian and Bibliographer, Iowa State University Library, Ames, Iowa, USA.