

9-1-2006

Image Database Task Force Final Report to the Board

Consortium of Academic and Research Libraries in Illinois

Follow this and additional works at: <http://digitalcommons.carli.illinois.edu/taskforce>

Recommended Citation

Consortium of Academic and Research Libraries in Illinois, "Image Database Task Force Final Report to the Board" (2006). *CARLI Task Forces*. Paper 5.
<http://digitalcommons.carli.illinois.edu/taskforce/5>

This Article is brought to you for free and open access by the CARLI Consortium Collections at digitalcommons@carli. It has been accepted for inclusion in CARLI Task Forces by an authorized administrator of digitalcommons@carli. For more information, please contact mchamber@uillinois.edu.

CREATING A DIGITAL IMAGE DATABASE
CARLI INVESTIGATIVE TASK FORCE REPORT
SUBMITTED ON SEPTEMBER 18, 2006

Task Force Members

Brandon Gant, CARLI

Jeff Hancks, Western Illinois University

Regina McBride, Southern Illinois University Edwardsville

Beth Sandore, University of Illinois at Urbana-Champaign

Task Force Co-Chairs

Gretchen Tuchel, University of Chicago

Suzanne Wilson, Illinois Wesleyan University

<i>Executive Summary</i>	3
<i>Phases of Implementation</i>	5
Phase I	5
Phase II	5
Phase III	5
Phase IV	6
<i>Architectural Sketch: Digital Object Management</i>	8
<i>Sources of Images</i>	9
Coverage Areas	9
Priority One Content	10
Content available for purchase	11
Content Available Freely on the Web	11
<i>Resources Required</i>	12
Staffing	12
Committees	12
Hardware/Software	13
<i>Metadata Production and Maintenance</i>	14
Harvested Content Metadata	14
Access to Non-OAI-compliant Visual Resource Collections	15
Metadata Maintenance	15
Collection Development Policy for OAI Metadata Harvesting	15
<i>Searching & Harvesting Rationale/Overview</i>	16
OAI Harvesting	16
Federated Search	18
Faceted Search	18
<i>Funding Models</i>	21
<i>Grant Opportunities</i>	22
<i>Authentication</i>	24
<i>Preservation</i>	24
<i>User Experience Defined</i>	24
<i>Conclusions</i>	27

Executive Summary

The CARLI Board of Directors established the CARLI Image Database Task Force in January of 2006. The committee was charged with developing a plan to build a database, in phases, of digital images to support learning, instruction, and research in all disciplines.

In its final phase, the database would include images from the following sources:

- Public images and/or metadata that are freely available on the Internet;
- Images that are purchased by or through CARLI for consortium members;
- Images that are created and/or owned by CARLI members.

Some of these collections will derive from CARLI's CONTENTdm implementation and others will come from different digital library collection systems. The Task Force investigations have revealed that CARLI institutions find value in aggregating access to visual resources to support research and learning. The real value comes in the creation of a common portal that brings together the access to these resources and makes them easily discoverable by users across the CARLI consortium. Other consortia have already made strides in this direction. The Task Force recommends that CARLI build on work that other consortia have accomplished, and extend it using proven technology to make easy and convenient the discovery and access to these critical resources.

Research for this report was produced through the investigations of the Task Force members who looked to similar efforts to aggregate visual resources and other content at institutions such as OhioLink, Wisconsin Library Services (WiLS), the University of Michigan's OAIster, and others. While these aggregations each have similar components, there is nothing equal in scope for a library consortium of this size on which we could base our recommendations. The biggest difference between the aforementioned projects and the one proposed in this report is that institutionally owned or created digital content was the focus of the aggregation. None of the other activities have attempted to aggregate free content with locally created as well as commercially available collections on behalf of a consortium for the purposes of supporting learning and research. This is a

new and worthwhile undertaking, and one that the Task Force believes is certainly possible, in a well-planned, multi-phased process.

The report outlines phases of implementation in addition to areas that need further development. The committee has described directions we may wish to pursue, but feels strongly that more development needs to be done in each area. This project is quite extensive in scope, and, depending on the availability of resources, this implementation could take three to five years.

In brief, the major recommendations included in the report are:

1. Fund the acquisition of additional commercial as well as freely available visual resources for CARLI;
2. Build a visual resource collection utilizing the recommended phases and technology framework in this report;
3. Create an oversight committee to guide recommended work groups to ensure a shared vision for the acquisition and utilization of this resource;
4. Allocate CARLI staff to develop and collaborate with CARLI institutions in this process.

Phases of Implementation

Phase I

This initial phase is already under way at CARLI, with the implementation of CONTENTdm in the consortium and also the purchase of the SASKIA Collection. After refinements to the system are implemented, and end-user training has been completed, collections will be added more quickly and by more institutions. Phase I deals only with CARLI CONTENTdm collections. While not an official part of this proposal, Phase I is a precursor to this project.

Phase II

Phase two is an undertaking already planned by CARLI. In this phase CARLI will begin work with the Illinois State Library and member institutions running their own CONTENTdm servers (there are ten of these instances). This phase creates a central metadata repository for the State of Illinois and CARLI. This component will expand access to include collections that are housed at the Illinois State Library (excluding government files) such as the Illinois Digital Archive. The purpose of this phase is to allow a single search across all CONTENTdm collections created within the State of Illinois.

Phase III

In this phase, all processes from Phase II will continue. The Task Force proposes creation of a Web portal combining search, aggregation, and discovery services that will provide organized and thematic access to digitized and born-digital collections from CARLI institutions and from other free and licensed resources. This access to significant scholarly, cultural heritage, and historical content would be enabled by use of the Open Archives Initiative (OAI) Protocol of Metadata Harvesting (OAI-PMH). The OAI Protocol provides a powerful mechanism to incorporate the hidden web of resources available through CARLI and other OAI-compliant institutions, including non-CONTENTdm collections, and to search across these disparate resources .

The committee has identified a list of “Sources of Images” (see **Appendix A**) that could be incorporated into the portal in phases.

This phase represents a transition from searching the multi-site server to searching the OAI aggregation, expanding from CONTENTdm alone to a broader group of digital collections. This includes collections created by CARLI members as well as OAI compliant content available on the Web. The consortium may need to devote resources to enable CARLI institutions to become OAI data providers. This will require training and limited use of technical resources at CARLI institutions.

There are a number of successful working examples of OAI metadata aggregations, such as OAIster at the University of Michigan, and the CIC Harvest Metadata portal. The University of Illinois could readily provide assistance to the CARLI consortium in setting up an OAI service.

Phase IV

In Phase IV, all functionality of Phase III will continue and the aggregation will begin to include non-OAI compliant metadata resources.

In order to locate and make use of these resources, the second of two internationally recognized technologies will be implemented: Federated search software using the SRU (search and retrieve using URL) protocol. Due to architectural or other design constraints, not all databases can export OAI-compliant metadata for aggregation. In these instances, federated search software uses the SRU (Search and Retrieve using URL) standard. SRU constructs a standard Web search query, submits it to the database, and enables the system to present the search results to the user along with the OAI aggregated metadata. In effect, two separate technologies (OAI and Federated Search) are in operation behind the scenes to provide the user with more comprehensive search results across a number of databases. The purpose of OAI harvesting combined with the federated search across collections is to improve access across selected resources.

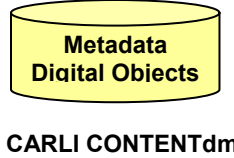
After CARLI purchases their own federated search tool, an interface can then be designed that would allow users to simultaneously search these different systems. A major digital

project already using federated searching along with locally produced content is the Wisconsin Library Services Project (WiLS).

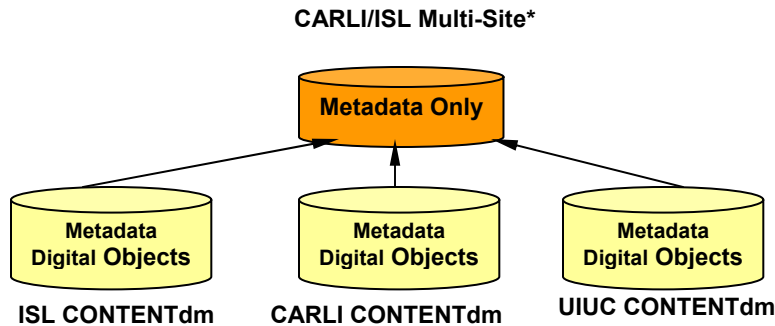
In addition to these tools, the committee also recommends that research be done into Faceted Searching. We feel that this tool could help enable thematic exploration of collections by providing more leads into content.

Architectural Sketch: Digital Object Management

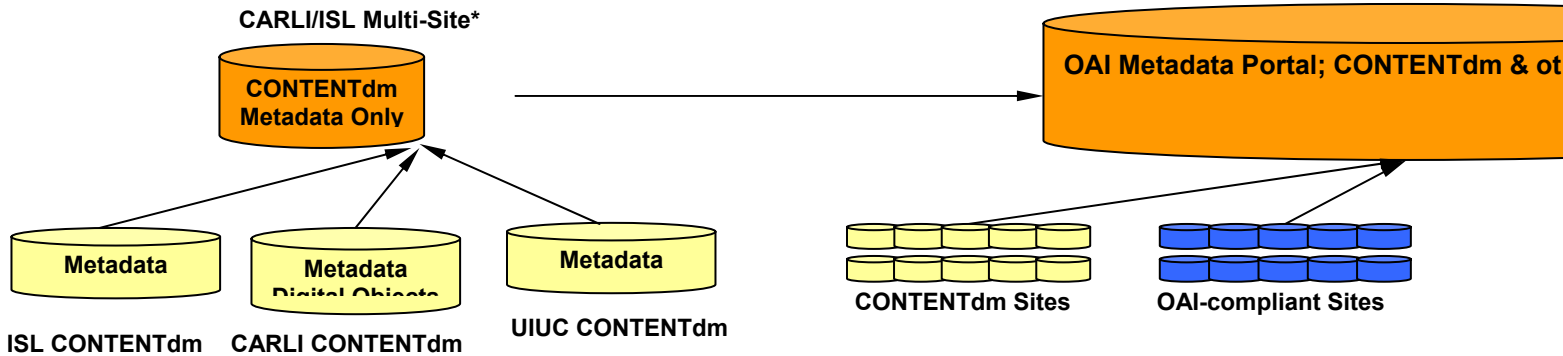
Phase I Allow CARLI members to create new collections on the CARLI Server
 Allow CARLI members to migrate existing collections to the CARLI Server



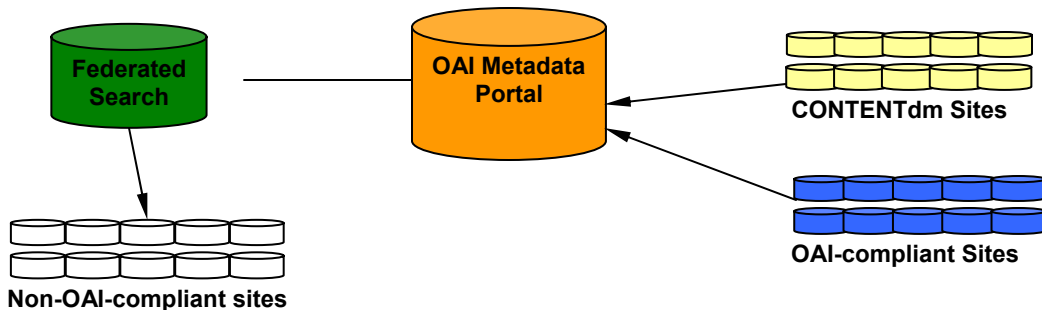
Phase II Setup CARLI/ISL (State Library) CONTENTdm Multi-Site Server
 This would be the central metadata repository for the State of Illinois



Phase III Implement OAI Provider/portal to aggregated metadata from free/licensed collections. Add/Link Digital Object Metadata to a Faceted Search engine



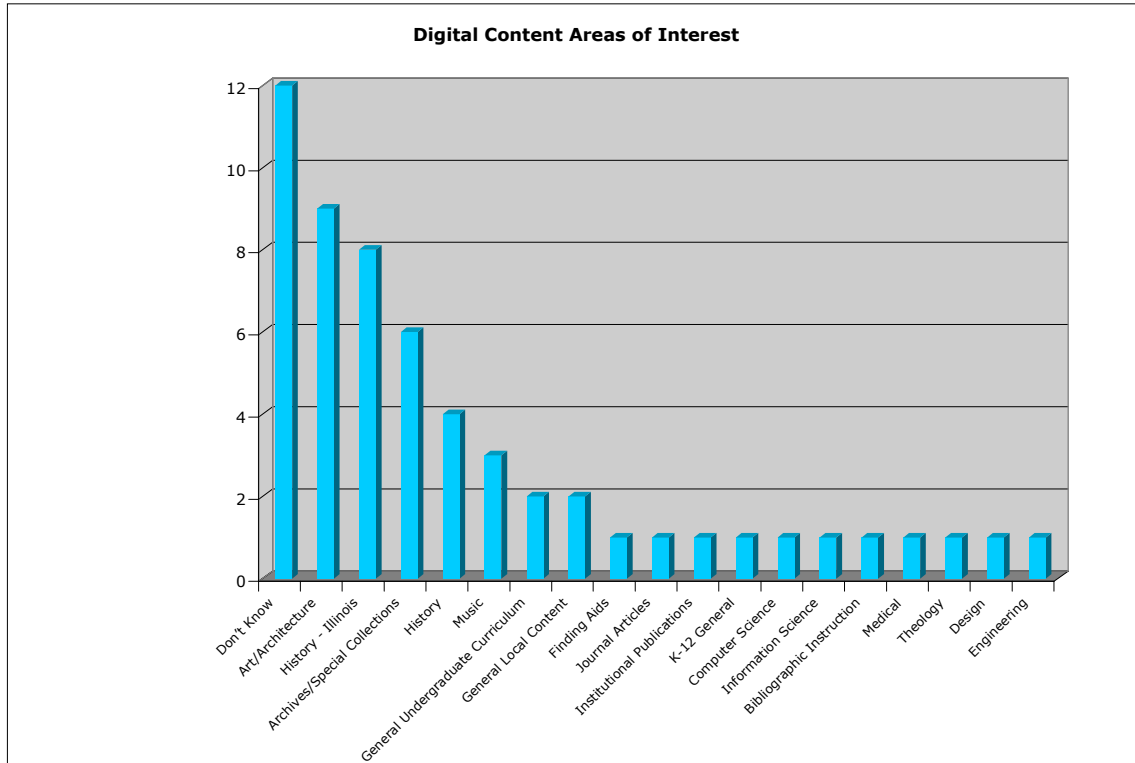
Phase IV Setup OAI or similar open standard metadata harvesting mechanism. Add non-CONTENTdm Metadata to the OAI metadata repository; use federated search to access data in non-OAI-compliant systems.



Sources of Images

Coverage Areas

Recommendations for images sources for topical areas are based on responses from the CARLI CONTENTdm Implementation survey conducted in 2005. Respondents' answers were categorized to develop broad areas of content. The following table depicts desired content coverage.



The largest number of responses indicated that either participants are not aware of what content might be available to them, or not aware of what content they might use at their institution. This demonstrates a need for outreach and education about what kind of content is available and how it can be used.

The next largest areas of interest were Art & Architecture, History (both Illinois and general), and digitized special collections from institutional collections. It is likely that many institutions' special collections will be contributors for Illinois historical images. Fortunately, much of the digital image content available freely or for purchase relates to cultural objects, those of art, architecture and history.

Based on the phased rollout of this proposal, sources of images have been prioritized in three stages. Content categorized as Priority 1 can be incorporated into the system during Phase III of the digital image database process. Content categorized as Priority 2 and/or 3 can be incorporated into the project in Phase IV.

Appendix A contains a spreadsheet of content. Each collection listed includes a description, Web address, OAI compliance (yes or no), and the Phase in which it is recommended for inclusion. While medical/scientific digital images were less requested in the survey, they are more readily available and easily integrated, and also fit more closely with the goal of this digital image database collection.

Priority One Content

Ideal content to add in Phase III is digital images with metadata that is free to CARLI. A prime example of this is Allan Kohl's Art Images for College Teaching.¹ "Art Images for College Teaching (AICT) is a personal, non-profit project of its author, art historian and visual resources curator Allan T. Kohl. AICT is intended primarily to disseminate images of art and architectural works in the public domain on a free-access, free-use basis to all levels of the educational community, as well as to the public at large." Because this collection has been photographed on location by Allan, who consents to their use in any application that is both educational and non-commercial in nature, copyright issues are not applicable. Allan has generously offered this collection to the project and the Task Force feels it would be an ideal addition, fitting desired content areas and as a test case to begin modifying metadata for OAI compliance. The collection can be acquired by borrowing the CDs from Mr. Kohl. The metadata is available in a spreadsheet format, and the only cost to the user is return postage.

Content that is freely available on the Web and is OAI compliant is recommended as priority two, or for Phase IV inclusion. An example of this is the American Memory project, available from the Library of Congress web site. The OAI harvestable portion of the collection includes images pertaining to world and U.S. historical photographs,

¹ <http://www.arthist.umn.edu/aict/html/>

architecture and interior design, art images, American sheet music, and digitized books, among others. Appendix A details further collections that are available for this Phase.

Content available for purchase

Content available for purchase is recommended for Phase IV inclusion. Vendor sites (listed in Appendix A) offer extensive art historical, architectural and scientific images. Scholars Resources, from whom CARLI licensed the Saskia collection of 30,000 art historical images, also has other collections that would be excellent additions to this project.

Content Available Freely on the Web

Web-based, free content will be the most difficult content to incorporate based on its varying degree of copyright restriction and metadata standards, both which affect harvesting capability. The approach to including this type of content is to allow the user to search it through a federated search interface, and then deliver the user to the images through hyperlinks. This is recommended for Phase IV.

Resources Required

The committee has identified several areas that will need to be addressed and will require either financial or time commitments. These are all critical to the successful implementation of a project of this magnitude.

Staffing

The committee feels strongly that a position for a Digital Library Developer will need to be created within CARLI that works specifically on the implementation and ongoing support of the image database. This person would help design and shepherd this project.

Minimum requirements should include:

- Previous experience and accomplishments working in digital repository management, electronic collections, etc.
- Experience creating and capturing a wide variety of digital formats.
- Working knowledge of metadata standards including Dublin Core, VRA, and MARC.
- Experience leading collaborative projects.
- Knowledge of the issues related to copyright, and intellectual access.
- Working understanding of faceted and federated search tools.

Committees

Due to the complexity of this project and the fact that this has not been attempted before as a consortium, there should be a Steering committee to guide and distribute the work being done by others to ensure consistency and maintain shared understanding. A first order of business would be to survey the membership to determine what digitization activities are being pursued or considered by other institutions in the consortium. The results of this survey would provide a starting point for these committees to focus their goals. Some of the committees we foresee being needed are:

Steering committee to oversee the following:

- Collection Development – Oversee the development of collections for different disciplines. Ensure that all areas are being represented adequately. Identify collaborative opportunities between institute members either digitizing or looking to obtain similar content.

- Database Management and Standards Committee – Develop guidelines for metadata to be used by contributors. This should include the OAI point person mentioned in the metadata section.
- Technology – Deals with issues related to metadata harvesting, federated searching, authentication, and faceted searching.
- Funding and Grant Opportunities—This committee would acquire content descriptions of collections that CARLI institutions are digitizing to match with available funding.
- Marketing and Communication--Includes training, user education, outreach, participation in web usability programming, and promotion of locally created collections.
- Web Design Committee—Establish standards for access to collections, create usable design based on institutional requirements.

Hardware/Software

The committee recommends that CARLI amend the WebFeat contract to purchase its own instance of the program and provide the tool for federated searching in Phase IV. Long-term planning should also include a resolution for a faceted search tool.

Also, after more research, if Faceted Searching is a desired tool a RFP process would need to be initiated to choose a specific product.

Metadata Production and Maintenance

Following the lead of OhioLink, the Task Force recommends the creation of a Database Management and Standards Committee. This committee will maintain quality standards for the central catalog by creating policies and procedures concerning all metadata for all members to follow. They will also work with the Technology committee to massage harvested data from free Web sites and to configure and map data from purchased sites. It is our recommendation that this committee consist of cataloging librarians; either those working in traditional libraries as catalogers or those who have worked as catalogers of digital materials such as images, audio or video recordings.

Member contributed collection metadata should be encouraged to follow a metadata standard of their choice, with a system component that will help in mapping metadata elements back to the Dublin Core application profile. It is expected that much of the data contributed to CONTENTdm will be created using varying degrees of standards. Among the projects consulted, Dublin Core has been the most widely used metadata standard. Our recommendation is to map data fields to the Dublin Core standard, allowing for the broadest use of data standards to be used, while maintaining common, searchable metadata. Providing technical advice and support for institutions interested in OAI metadata should be supported. Ideally, all collections that are to be part of CONTENTdm and therefore the digital image database, should adhere to the *Best Practices for OAI Data Provider Implementations and Shareable Metadata* document available at:

<http://oai-best.comm.nsdlib.org/cgi-bin/wiki.pl?TableOfContents>

Harvested Content Metadata

Harvesting metadata from OAI compliant collections, as opposed to federated searching of the sites, will provide a richer end-user experience through the massaging of data and presentation through one user interface. **Appendix A** includes indicators of sites that are OAI compliant. This is the ideal method for creating the digital image collection. Metadata harvested through OAI means can be in any format as long as a simple Dublin core record is available for each item.

Access to Non-OAI-compliant Visual Resource Collections

Content that cannot be harvested using the OAI protocol certainly provides a suitable instance for federated searching. WebFeat is an example of a federated search tool that can be used to locate this content and mirror it to end users for selection.

Metadata Maintenance

Central coordination for the reporting of results and observations and day-to-day operation and maintenance of harvesting services is recommended. Typical staff time required to implement OAI-PMH is at most 320 hours assuming accessible metadata with defined Dublin Core mapping. This includes initial development time, programming and customizing Open Source tools, writing supplemental CGI scripts, creating XSLT or other metadata transforming utilities. Graduate assistants or other part-time programming staff may do the bulk of this work. A small amount of ongoing maintenance is required for dealing with errors in metadata, ongoing modifications in metadata workflow, character encoding issues, and responding to bug reports from harvesting agents. This work could be incorporated into a system administration workflow.²

Collection Development Policy for OAI Metadata Harvesting

The task force recommends implementing the policy used by OAIster.

University of Michigan OAIster Collection Development Policy:

- We harvest and retain all records that point to digital objects.
- This includes freely-available and restricted-access digital objects.
- For those repositories that have a small number of records with incorrect UTF-8 or XML that causes our transformation engine to fail, we will fix these records such that the engine can successfully complete. However, because we have to handle these repositories more than once each time we harvest and consequently cause time delays, they are harvested on a monthly basis. Those repositories containing many incorrect UTF-8 or XML records will be dropped. We will communicate these problems with the repository owner as time allows.
Repositories with no UTF-8 or XML errors are harvested on a weekly basis.

² Found on July 31, 2006 at:
<http://cic harvest.grainger.uiuc.edu/cicproposal.asp>

- When harvesting on a regular basis (weekly or monthly) has failed at least three times and we are unable to discover a new OAI based URL for a repository, we may drop the repository. As above, we communicate these problems with the repository owner as time allows.
- For repositories with fewer than 5 records that point to digital objects, we hold off on using this repository. Communication by the repository owner of an increase in records is the most efficient method to re-engage our harvesting of the repository.³

Searching & Harvesting Rationale/Overview

We seek to include many rich and dispersed visual resources in this framework, but not all are accessible using one search mechanism because they have been developed using various search protocols and there is no standard search protocol for digital content on the Internet. Within the past several years, two approaches have emerged that allow a degree of standard searching across these many different databases. These approaches are Open Archives Initiative Protocol Metadata Harvesting (OAI-PMH) and Federated Search. More recently, a new approach--faceted searching--has been developed. In this section we present an overview and a rationale for the ways in which the CARLI community can gain access to disparate visual resources using these search technologies.

OAI Harvesting

Open Archives Initiative Protocol Metadata Harvesting (OAI-PMH) is designed to enable resource discovery across distributed and heterogeneous collections. Originally developed to facilitate interoperability among e-print archives, OAI-PMH is now in use by numerous communities to expose and allow aggregation of metadata describing a wide range of collections.

Numerous institutions have played pioneering roles in the design and development of the OAI-PMH. The University of Illinois and the University of Michigan were among the first to establish OAI-compliant metadata harvesting services, with funding provided by the Andrew W. Mellon Foundation, and more recently from the CIC. CARLI could take

³ Found on July 21, 2006 at:
<http://oaister.umdl.umich.edu/o/oaister/>

advantage of this work as they consider approaches to providing access to OAI metadata aggregations. A search of the CIC OAI-Harvest metadata portal yields both text and thumbnail images from collections at the University of Illinois.

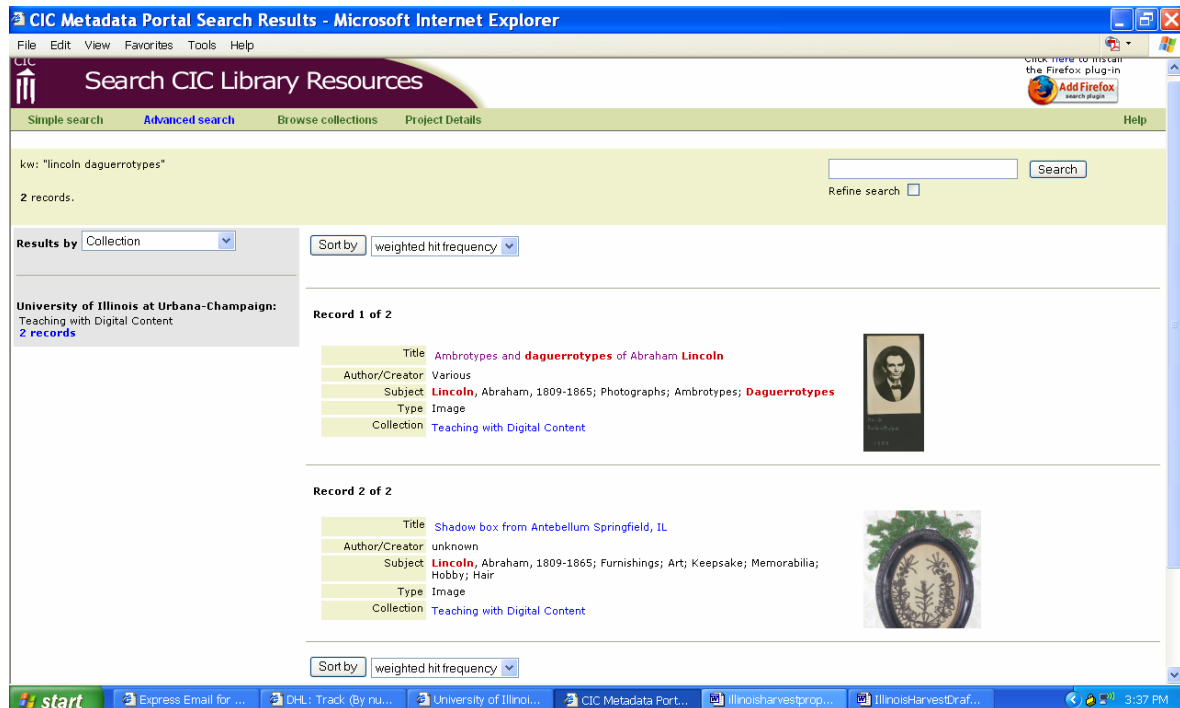


Figure 1: A sample search of the CIC OAI Harvest metadata portal for “Lincoln daguerrotypes” yields materials from the University of Illinois Library.

The University of Michigan’s OAIster service utilizes the same underlying technology to provide access to over 8.7 million digital objects held in collections across hundreds of institutions worldwide.

The screenshot shows the OAIster search results page. At the top, the OAIster logo is displayed with the tagline "...find the pearls". A navigation bar includes links for Home, Search, Browse Institutions, and Help. A "View Bookbag" link is also present. The search results summary indicates that the search was performed in the "entire record" field for the terms "american", "indian", and "dance". It reports finding 136 records and provides options to revise the search or view the first 10 records.

The "Results by Institution" section lists several sources and their respective record counts:

- American Museum of Natural History (AMNH) Research Library: 8 records
- Archive of Popular American Music (APAM), University of California Los Angeles (UCLA) Music Library: 2 records
- California Digital Library (CDL): 3 records
- Capturing Electronic Publications (CEP) Archive: 4 records
- Colorado Plateau Digital Archives, Cline Library, Northern Arizona University (NAU): 1 record
- Digital Manuscripts, Archives, and Special Collections, Washington State University (WSU) Libraries: 34 records
- HighWire Press, Stanford University: 1 record

The "Record 1 of 136" details are as follows:

Sort by	title
Record 1 of 136	
add to bookbag	
Title	; ; Parson Motanic, Cayuse Indian , in costume. ; Parson Motanic, Cayuse Indian , in costume. ;
Author/Creator	Moorhouse, Lee; nsg
Contributor	nsg
Publisher	Joint Project of the University of Oregon Libraries and the Tamástslikt Cultural Institute of the Confederated Tribes of the Umatilla
Year	2003-07-25
Resource Type	Image
Resource Type	Image
Resource Type	Outside Photograph;
Resource Format	image/jpg ,295,352 bytes,Scanned in transmissive mode with Silverfast AI 6.0 on UMAX PowerLook III flatbed scanner: Levels adjusted from 0 - 1.00 - 255 to 14 - 1.00 - 134 prior to capture, 125 dpi ,Scanned images enter Photoshop with negative polarity and have a reversed orientation. Archival Tiffs are saved. Adobe Photoshop used to reorient and crop images, invert image to positive polarity, set and neutralize shadow and highlight point, adjust midtone contrast, and desaturate color information. Second production tiff saved, retaining adjustments layers. Photoshop used to convert color profile from Adobe 1998, to sRGB, set resolution to 125 dpi and resize the largest pixel dimension of the image to 875. Display jpeg saved after resizing.

Figure 2. Results of an OAIster search for images on the topic of “American Indian Dance” reveal images in collections ranging from the California Digital Library to the National Library of Australia.

Federated Search

Federated search technologies provide simultaneous search capabilities over the myriad of distributed information resources. They “broadcast” the user search terms to multiple databases, sources, and platforms all at one time and collect search results for display of information and links to users. These systems have numerous benefits, including: (1) they can be tailored to user needs; (2) they provide access to resources through the federated search view and they connect the user directly to the native interface where the content resides; and (3) the user search argument is translated into the terms accepted by the native search interface.

Faceted Search

Faceted searching is a new technology that should be explored to provide deeper, thematic exploration of these visual collections. As a search interface for OAI compliant data, this would benefit the user by providing more leads into content.

"A faceted classification system allows the assignment of multiple classifications to an object, allowing searching and browsing of related information through several classes. Elements may include subject, geographical, temporal and form

of an item. In an online world, faceted classification becomes much more powerful than more traditional hierarchical schemes such as the Library of Congress classification because it provides multiple navigational paths to any one item of information.”⁴

For instance, an image depicting an archaeological site could be classified by location, date of dig, culture, artifact types recovered, archaeologist, and time period. A user could navigate through any of these facets, combining them in any way to reach exactly the desired results. A good example of faceted search implementation is OCLC’s Worldcat catalog Beta product, which provides the user options to refine a search by providing choices from a variety of subject categories. The following screen shots demonstrate how a basic search can be explored in many different categories.

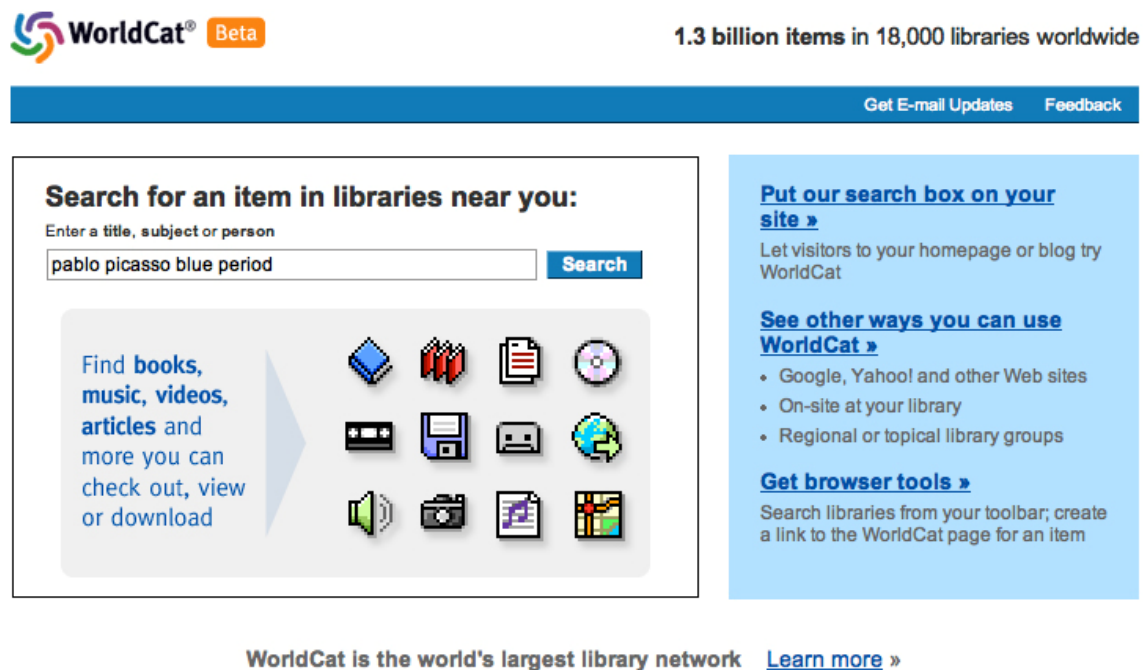


Figure 3. WorldCat search using basic search terms.

⁴ Found on September 13, 2006 at: http://en.wikipedia.org/wiki/Faceted_classification

Search results for 'pablo picasso blue period'









Refine Your Search	Results 1-10 of about 31 (.08 seconds) << First < Prev 1 2 3 Next >
<p>Author</p> <p>Pablo Picasso (9)</p> <p>Temma Kaplan (3)</p> <p>Sidney Janis Gallery (3)</p> <p>Perry Miller Adato (1)</p> <p>Kevin J Anderson (1)</p> <p>Andrew Langley (1)</p> <p>Content</p> <p>Art & Architecture (15)</p> <p>Sociology (3)</p> <p>Language, Linguistic... (2)</p> <p>Music (2)</p> <p>Format</p> <p>Book (18)</p> <p>Visual Material (9)</p> <p>Musical Score (2)</p> <p>Internet Resource (2)</p> <p>Language</p> <p>English (27)</p> <p>Spanish (2)</p> <p>Year</p> <p>1960 (4)</p> <p>2005 (3)</p> <p>2004 (2)</p> <p>1992 (2)</p> <p>1900 (2)</p> <p>1986 (2)</p>	<ol style="list-style-type: none"> War of the worlds : global dispatches by Kevin J Anderson Type: English :  Book : Fiction Publisher: New York : Bantam Books, 1996. Pablo Picasso by Leonie Bennett Type: English :  Book : Primary school Publisher: Chicago, Ill. : Heinemann Library, ©2005. Paintings, blue period by American Library Color Slide Company. Type: English :  Visual Material : Slide Publisher: New York : American Library Color Slide Co., [197-?] Picasso : masterpieces of the blue period : 38 paintings by Pablo Picasso; Laszlo Glozer Type: English :  Book Publisher: New York : W.W. Norton, [1992?], ©1988. Picasso, 1881- : his Blue Period (1900-1905) : pastels, water-colours, and drawings : also the complete set of small bronzes of female figures, 1945-47. by Pablo Picasso; Sidney Janis Gallery. Type: English :  Book :  Microform Publisher: New York : Sidney Janis Gallery, 1960. Picasso 1881-, his blue period : collection of pastels, water-colours, and drawings 1900-1905, also the complete set of small bronzes of female figures, 1945-47. by Sidney Janis Gallery. Type: English :  Book :  Microform Publisher: London : The Gallery, [1960]

Figure 4. WorldCat showing expanded search categories offered in five different areas.

The combination of these tools offers a way to unlock the hidden web of resources that are available across the CARLI consortium (e.g., local databases, commercially-licensed resources) and to make them widely accessible to the CARLI membership. The strength of combining a variety of search and harvest approaches will provide a single search point for CARLI users to obtain access to freely available as well as licensed visual resources.

Funding Models

The members of the CARLI Imaging Databases Task Force recognize the challenges and complexities of funding for major projects and initiatives in a large and diverse consortium. Facing similar questions regarding membership and equitable financial models, as recently encountered by the CARLI Board, the Task Force explored several possibilities and options for member libraries.

After conversations with administrators in other large academic library consortia such as OhioLink and drawing upon successful past practices in Illinois, the Task Force developed several key components for funding digital imaging databases. The recommended model is as follows:

- CARLI will function as an aggregator for all free collections. Free collections would be available to all CARLI member libraries based on the algorithm presently in place for CARLI membership options.
- CARLI will also negotiate attractive pricing options for member libraries to purchase databases. These purchased databases will be centrally housed and made available to the purchasing library, or to a group of CARLI libraries that may choose to share in a joint purchase.
- CARLI member libraries choosing to take advantage of this service would “pay as they go.” For the Consortium this would be a value neutral proposition

Grant Opportunities

Listed below are foundations and agencies that have funded imaging projects in the past and/or have goals that include funding library initiatives in digitization and preservation. They may provide additional opportunities for CARLI libraries.

Andrew W. Mellon Foundation

<http://www.mellon.org>

The Mellon Foundation supports projects in the following areas: higher education and scholarship, scholarly communication, research in information technology, museums and art conservation, performing arts and conservation and the environment. Awards are generally in excess of \$100,000.

Annenberg Foundation

http://www.annenbergfoundation.org/grants/grants_show.htm?doc_id=209659

The Annenberg Foundation provides support for projects within its grant-making interest areas of education and youth, arts and culture, civic and community and health. During the past three years the Foundation has funded nearly 50 university proposals. Funding to libraries and universities are frequently in excess of \$100,000.

Gladys Krieble Delmas Foundation, Research Library Program

http://www.delmas.org/programs/research_lib_d.html

The Foundation encourages humanistic inquiry and artistic creativity by supporting research libraries and other institutions dedicated to our cultural heritage. Under the Research Library Program the Foundation has awarded grants under \$60,000.

Institute of Museums and Library Services

National Awards for Museum and Library Service

National Leadership Grants

<http://www.imls.gov/applicants/name.shtm>

The Institute offers several grant programs that provide opportunities to museums, libraries and library agencies. These programs are designed to further collaborative digital projects such as, developing digitized collections. Funding to libraries range from \$50,000–\$1,000,000.

Library Services and Technology Act Grant

<http://www.cyberdriveillinois.com/departments/library/>

LSTA funding is provided by a grant from the Institute of Museum and Library Services to the Illinois State Library under the provisions of the Library Services and Technology Act. Digital Imaging is an area designated by the Illinois State Library, as a high priority for LSTA funds. Library collaborative projects are well received. Funding ranges from a few thousand dollars to over \$100,000.

Below are two agencies that may support imaging projects in specialized areas.

National Endowment for the Humanities, Preservation and Access

<http://www.neh.gov/GRANTS/grantsbydivision.html#public>

NEH supports initiatives to create intellectual access to collections that are considered important for research, education, and public programming in the humanities. Collections may include, but are not limited to, books, journals, newspapers, manuscript and archival materials, maps and still and moving images. Funding levels for awards range from \$50,000 to a maximum of \$350,000.

Siragusa Foundation (Midwest region, special interest in Chicago)

<http://www.siragusa.org/pages/guidelines/14.php>

Culture and education are two of the five areas funded by the Foundation. Grants are awarded primarily to institutions in the Chicago area, however other areas of Illinois such as, Bloomington have received money. A few other organizations across the country have also received support. Awards generally range from a few thousand dollars to approximately \$30,000.

Authentication

Access to consortial content can be presently managed through IP based recognition. The task force recommends looking to other projects future experience with new technologies (such as OhioLink's use of Shibboleth). It is in the best interest of the project to implement a more fine-grained authentication and authorization process.

Preservation

As with other CARLI digital resources, it is important to begin digital preservation of content created by CARLI institutions as well as preservation of digital content purchased and owned outright by CARLI (e.g., Saskia Scholars Resource). This will require building a consortial understanding of the importance of long-term digital preservation of content purchased outright with CARLI funds. Further, knowing that not all institutions have the local resources to support preservation of their born-digital or digitized content, the CARLI consortium ought to consider what, if any role it will play in long-term preservation of content that its members make available through the consortium's portal and search/discovery mechanisms.

User Experience Defined

The committee spent time researching what other institutions have done with their user interfaces' for digital library collections. At the time of these discussions we were envisioning one interface that all libraries would be able to access. This is not to limit the ability of libraries to one day "customize" this interface, but we felt this was a good starting point for everyone.

Two elements that we feel are critical to having a useful interface for the user are:

- 1.) The ability to easily search across collections regardless of their origin.
- 2.) Organized and thematic access to digitized and born digital collections

Two tools that could be very useful in implementing these elements are a federated search tool and a faceted search tool. Both of these should be explored further to examine what the advantages are of each of these systems.

The interface the task force most preferred is the Library of Congress' "American Memory." In the screen shot below you can see that the user is able to search for a specific object across all collections or the user may browse the collections by topic.

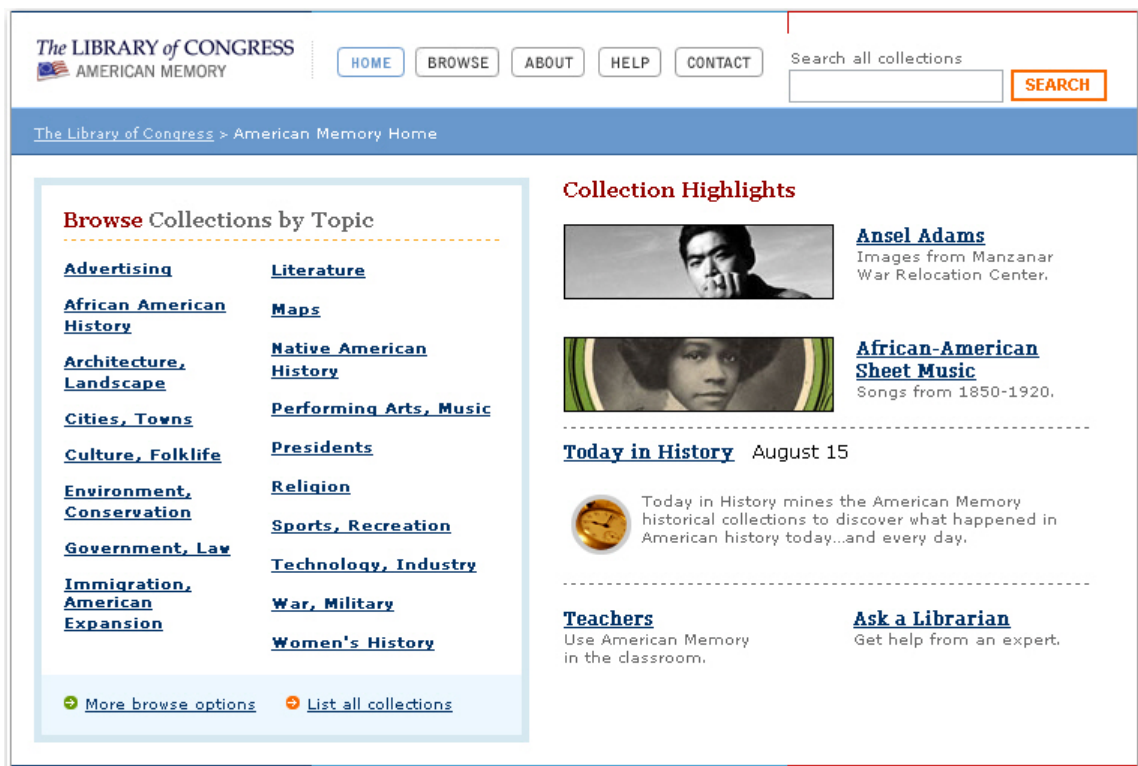


Figure 5. Library of Congress, American Memory

They also have an additional page with more options that allows you to browse by time period, format, etc. Thoughtfully organized browse options are very useful to the user who may be unsure about what they are looking for, but who can navigate to desired content through broad thematic categories.

Browse by Topic

- Advertising**
[Early Advertising, more](#)
- African American History**
[Slave Narratives, more](#)
- Architecture, Landscape**
[Historic Buildings, more](#)
- Cities, Towns**
[New York City Films, more](#)
- Culture, Folklife**
[September 11, Dust Bowl, more](#)
- Environment, Conservation**
[Florida Everglades, more](#)
- Government, Law**
[Continental Congress, more](#)
- Immigration, American Expansion**
[Chinese in California, more](#)
- Literature**
[Walt Whitman, more](#)
- Maps**
[Civil War Maps, more](#)
- Native American History**
[Pacific Northwest Culture, more](#)
- Performing Arts, Music**
[Music for the Nation, Creative Americans, more](#)
- Presidents**
[Washington, Jefferson, Lincoln, T. Roosevelt, more](#)
- Religion**
[Southern Black Churches, more](#)
- Sports, Recreation**
[Jackie Robinson, more](#)
- Technology, Industry**
[Alexander Graham Bell, more](#)
- War, Military**
[Civil War Photographs, WWI Rotogravures, more](#)
- Women's History**
[National Woman's Party Photographs, more](#)

List all collections

Browse Collections by Time Period

- [1400-1699](#) [1700-1799](#)
- [1800-1849](#) [1850-1899](#)
- [1900-1929](#) [1930-1949](#)
- [1950-1969](#) [1970-present](#)

Browse Collections Containing

- [Maps](#)
- [Manuscripts](#)
- [Motion Pictures](#)
- [Sheet Music, Song Sheets](#)
- [Photos, Prints](#)
- [Sound Recordings](#)
- [Books, Other Printed Texts](#)

Browse Collections by Place

- [Northeast U.S.](#) [All U.S.](#)
- [South U.S.](#) [International](#)
- [Midwest U.S.](#)
- [West U.S.](#)

Figure 6. Library of Congress, American Memory browse categories.

The Collaborative Digitization Program (<http://www.cdphheritage.org>) site was favored as an informational resource for its “Teacher Toolbox” and “Digitization Toolbox” that provide participants links to scanning standards, best practices, and more.

collaborative digitization program

[cdp home page](#)
[about cdp](#)
[get cdp info](#)
[contact cdp](#)

[Digital Toolbox](#) [Teacher Toolbox](#) [Collections](#) [Exhibits](#) [Projects](#) [Membership](#)

Search

Newspaper Collection
Heritage West
This site

Teacher Toolbox Menu

- [Manual for Program Development](#)
- [Discovering Digital Treasures](#)
- [Amache Through Primary Resources](#)
- [Model Lesson Format](#)
- [Planning Professional Development](#)
- [About Teaching with Colorado's Grant](#)
- [Teaching with Colorado's Historic Newspaper Collection](#)
- [Resources](#)

Teacher Toolbox

Why Primary Resources

Using primary source materials can help develop critical thinking and research skills, for learners of all ages. Used as complements to traditional methods, these resources can also enrich your classroom! This section provides resources and information for educators who are just getting started using digitized primary source materials, as well as those who are looking to broaden their skills or those of their colleagues.

Lesson Plans

Ranging from alpine plants o Works Progress Administration programs in Colorado, these lesson can be used as-is or as a good starting point to using digitized primary sources materials with students.

[view all lessons](#)

Professional Development

Self paced courses and other resources that offer a first experience or a brush up on searching for and utilizing photographs, maps audio and more!

[learn more](#)

Figure 7. Colorado Collaborative Digitization Program

An example of an OAI metadata aggregation known as OAIster (<http://oaister.umdl.umich.edu>) at the University of Michigan demonstrates the useful feature of searching metadata of various collections regardless of their origin and then displaying the object if requested. The downside of this interface is that there is no ability to browse by category.

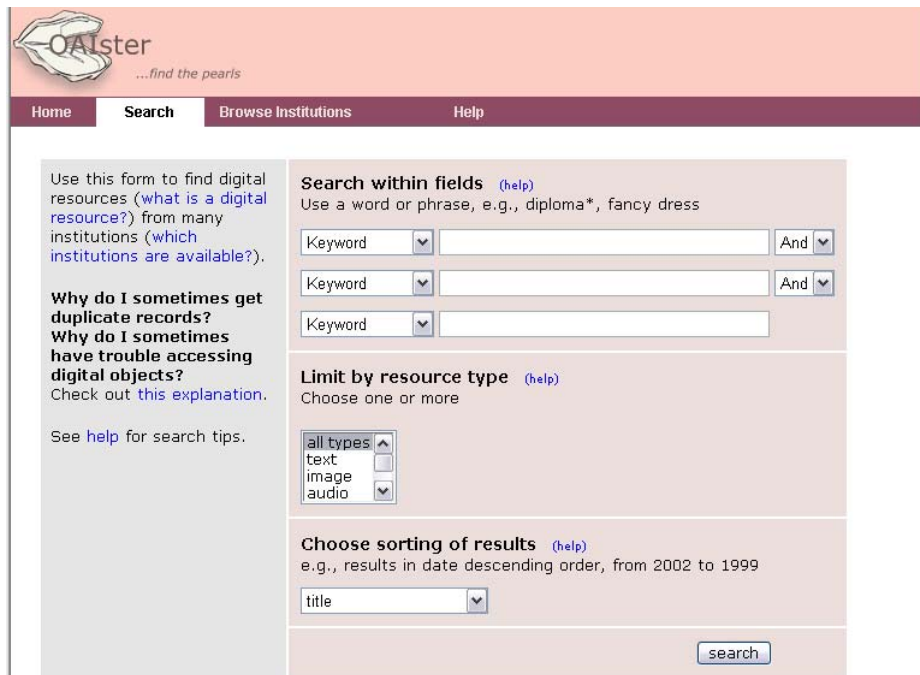


Figure 8. OAIster search interface.

Conclusions

We envision the CARLI digital image database framework not as a standalone product, but rather as a set of resources that can be combined with teaching and research tools across CARLI institutions.

The task force feels there may be an opportunity to communicate with other CARLI committees and task forces to see if there are some collaboration opportunities. Committees that appear to have a natural tie-in to the focus of our task force are the Learning Object Task Force and the Library Tools Task Force.

Looking beyond a digital image database, to the incorporation of digital images with a more broadly based research tool, CARLI might look at the SPIDER project (Shared Pedagogical Initiative: A Database of Electronic Resources for the University of California Community), which “experimented with the integration of resources for teaching, writing and research to undergraduates.” A project lead for SPIDER reported that students using this repository produced better writing, used the library more frequently, and conducted more research. Incorporating this digital image database with other types of learning/research tools could produce an even richer educational experience for faculty, students and the public in general.