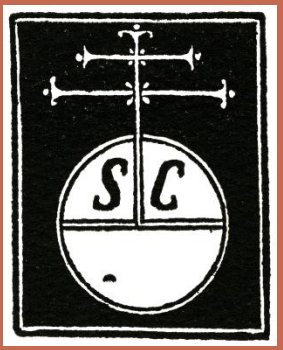
FROM W. E. B. TO WEB

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011



100K ITEMS PAPER TO WEB

Special Collections and University Archives (SCUA) at the W. E. B. Du Bois Library, UMass Amherst has embarked on a journey to bring the entire manuscript collection of sociologist, historian, social activist and author, W. E. B. Du Bois into the 21st century. This digitization project will produce a free, online interface that will allow unlimited access to over 150,000 items of great historical importance by keyword or selected field.

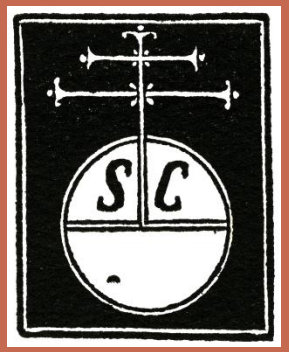
The project team has developed a sound process for digitization and description. A best practices document for describing manuscripts using MODS was created to support the unique demands of the collection. As of June 2011, the mid-point of original funding from the Verizon Foundation, the project team has scanned over 130,000 items and created 41,000 metadata records. With additional support from the National Endowment for the Humanities, SCUA will be able to make the entire digital collection available online by the end of 2013.

Our poster illustrates the project work flow from re-boxing and scanning to presentation on the web in our digital collection interface. This diagram includes item examples, job descriptions, implemented standards, samples of MODS templates, quality control methods, and screen shots of the online digital interface.

We are happy to share our challenges and success with our ALA colleagues.
Thank you for the opportunity.

Jeremy Smith, M.S., Digital Project Manager, W. E. B. Du Bois Digitization Project
Special Collections and University Archives at the University of Massachusetts
Abigail S. Baines, M.S., Electronic Reserves, Du Bois Library, & former Metadata Specialist for the
W. E. B. Du Bois Digitization Project, University of Massachusetts

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of W. E. B. Du Bois
Poster Session - ALA 2011



ABOUT CREDO

Credo is an online repository containing the digital collections held by the UMass Amherst Libraries' Department of Special Collections and University Archives (SCUA). SCUA presently houses more than 30,000 linear feet of archives and rare books, the vast majority of which are available only to users who are able to visit our collections in person. In an effort to make these unique materials more widely and freely accessible, SCUA and the Libraries' Systems Department began to develop Credo in 2009, with the goal of making the repository a central hub for exploring SCUA's collections from anywhere at any time.

With the generous support of the Verizon Foundation and the National Endowment for the Humanities, SCUA simultaneously embarked upon the massive project of digitizing the complete papers of the African American intellectual and activist, W.E.B. Du Bois. By the time that Credo was launched in June 2011, more than 40,000 digital objects from the Du Bois Papers were already available, and within two years, we expect to include the remainder of the collection -- another 60,000 items. SCUA plans to add significant content to Credo in the coming years both digitized and born digital, representing the breadth of our collecting interests from university history to social change, the history of New England, and innovation and entrepreneurship.

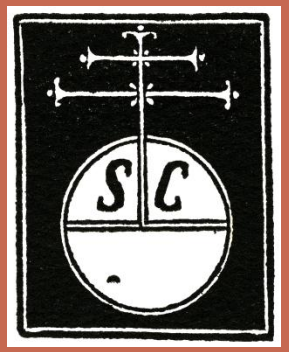
We do not expect the work of Credo ever to be truly finished and as such we welcome comments and feedback from our users. To send us your comments, please email us at credo@library.umass.edu.

Naming Our Digital Repository

In considering a name for our digital repository, we wanted to select something that would hold meaning for a variety of collections, but that would also pay homage to W.E.B. Du Bois, our digital pioneer. First published in 1904, *Credo* is a powerful statement of Du Bois's philosophy for achieving racial equality, standing in contrast to the approach taken by many other civil rights activists at the time. Credo speaks to peace, liberty, education, and patience, qualities that run throughout our many social change collections.

To learn more about *Credo* and to read the full-text, visit DuBoisopedia.

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of W. E. B. Du Bois
Poster Session - ALA 2011



DIGITIZATION PROJECT

FACTS

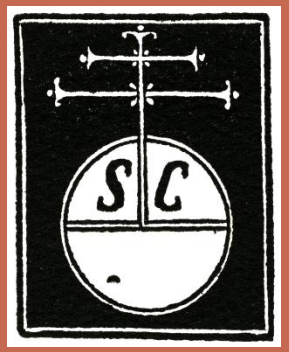
FUNDING
Exclusively from:
The Verizon Foundation and The National Endowment for the Humanities.

PERSONNEL

Project Director:	1 FTE
Metadata Specialists:	2 FTE
Scanners:	3.5 FTE (5-10 part-time positions, primarily students)

- The entire collection is being described at the item level
- There are approximately 150,000 items in the collection
- Dates of the collection range from the 1870s to the 1960s
- Portions of the collection were microfilmed in the 1980s
- As of the June 2011, 41,000 metadata records have been created and 130,000 images scanned.

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011

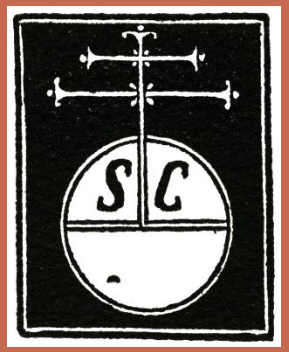


DIGITIZATION PROJECT

WORKFLOW

1. Original materials are re-foldered and re-boxed
2. Items are scanned as 300 dpi TIFF files and given a unique filename, e.g., mums312-b034-i393-001
3. Metadata records are created in the oXygen XML editor and encoded using the Metadata Object Description Schema (MODS). Local best practices for descriptive metadata have been developed for the item-level cataloging of manuscripts. Implemented standards include: AACR2, AAT, DACS, LCSH, plus local name and subject authorities.
4. Digital objects are checked for errors by both the Metadata Specialists and the Project Manager. The Project Manager uses Python scripts to analyze both the image scans and the metadata records.
5. MODS records are wrapped in the Metadata Encoding and Transmission Standard (METS). There is a METS record for each digital object. The METS record also defines the digital objects by describing which image files make up a single object and stores technical metadata about the individual image files. All files are then ingested into a digital repository based on Fedora, the open source digital repository framework.
6. A user interface, ***Credo***, will provide access to the entire Du Bois collection, as well as all of SCUA's digital materials. The interface features item and collection level description, simple and advanced search, and faceted browsing

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011



credo

Special Collections & University Archives

home | browse | about

search digital collections

advanced search

W.E.B. Du Bois Papers

Search and view correspondence, writings, and photographs in the W.E.B. Du Bois Papers. more>

Du Bois Photographs

The W. E. B. Du Bois Papers contains nearly 900 photographs spanning Du Bois's entire lifetime; explore the collection's images. more>

Credo

Welcome to Credo, the home of SCUA's digital collections. Learn more about Credo and what you can find inside. more>

Special Collections and University Archives: W.E.B. Du Bois Library: University of Massachusetts Amherst | contact us

DU BOIS ONLINE

"I BELIEVE IN LIBERTY FOR ALL MEN"

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011

HOW TO USE CREDO

We have built Credo so that finding what you want is easy, but it doesn't hurt to share a few tips to make it even easier.

If available, all of our items are cataloged with the creator, date of creation, the location of creation, Library of Congress subject headings, and brief natural language abstracts. When using our **basic search** you will get results from all of these fields. If you wish to find something from only one of these fields, or from specific combinations of fields, we suggest using the **advanced search**.

Basic Search

Just like Google, the search box is a good place to start for general inquires. Searches done here will search all keywords, phrases, formats, fields, and collections. For example, if you type 'Marcus Garvey Pan African Congress', the search tool will search every collection in Credo and locate every item in which those words appear together. To narrow your search to an exact word or phrase you can use quotes.

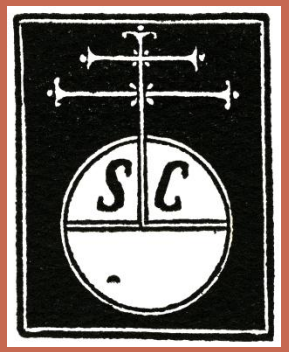
Advanced Search

The advanced search is where to go for more specific inquiries. The advanced search will allow you to limit your search to specific fields such as name, date, place, title, or abstract and will allow you to combine searches with Boolean operators such as "and", "or" and "not'. If you know the item identification number, you can enter it here as well.

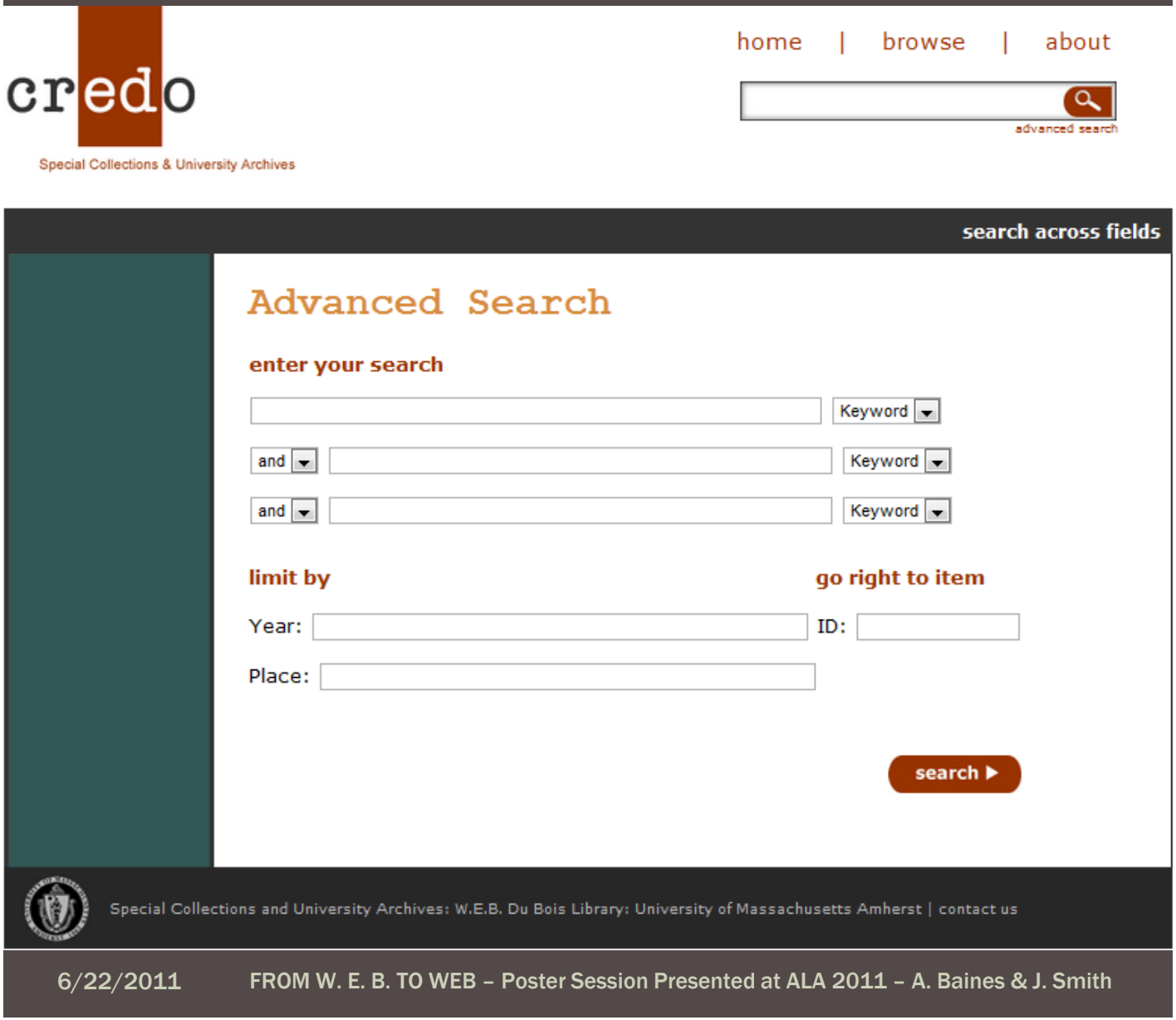
Please note that search terms limited to a particular field will not give you every instance of that term in Credo. For example, if you search for 'Du Bois Literary Prize' only in the subject field, your results will not return mentions of the prize that appear only in the Abstract. Usually if a name, topic, or place is listed in the subject field, it constitutes a significant portion of the intellectual content of the item, but if you want *every* mention of a word, it is best to do a keyword search.

Once you find what you want on the results page, you can click on the title to see the item view. The item view will give you all of the information about that item (aka the metadata) and the ability to examine the digitized item.

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of W. E. B. Du Bois
Poster Session - ALA 2011



HOW TO USE CREDO



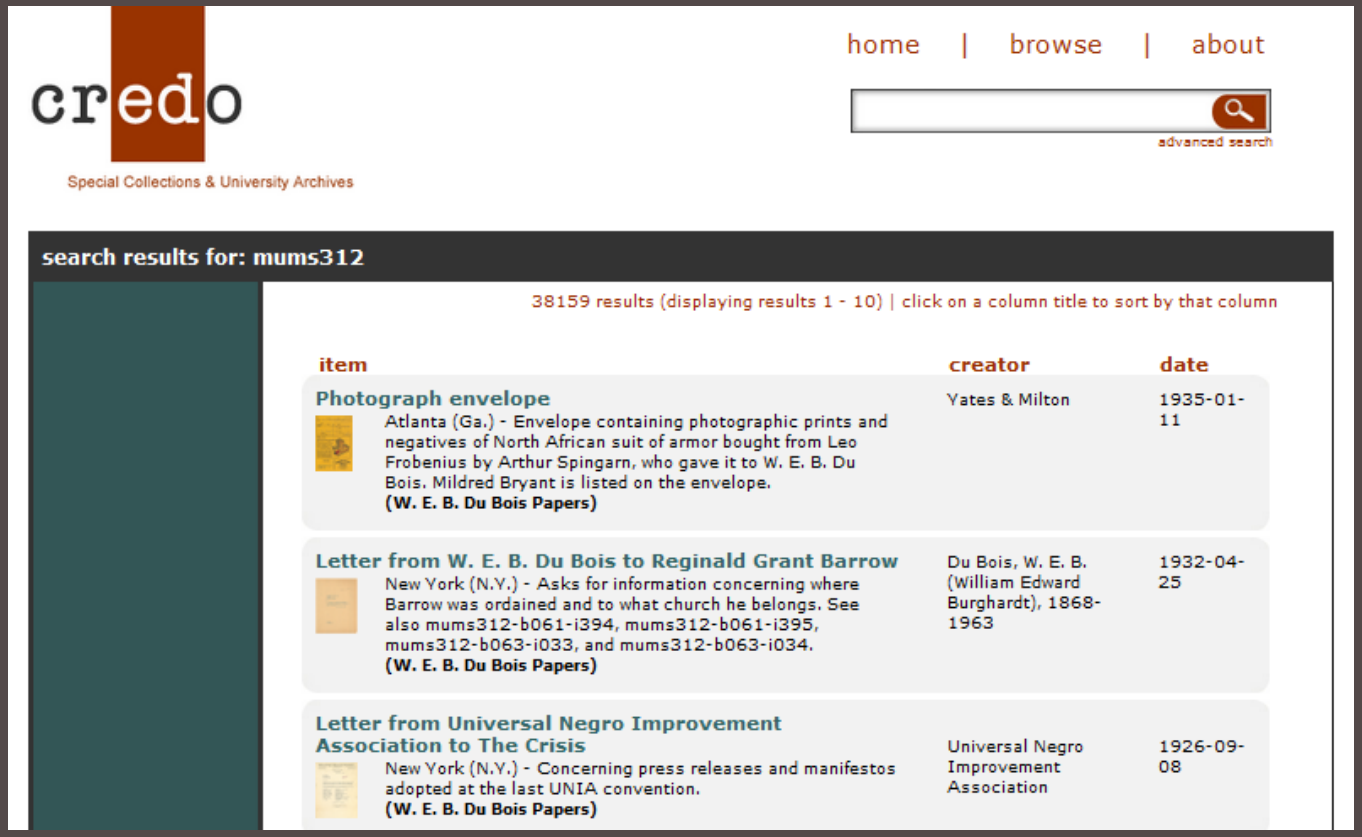
DU BOIS ONLINE
"I BELIEVE IN LIBERTY FOR ALL MEN"

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011

HOW TO USE CREDO

Collections in Credo

Each digital object that you view in Credo is part of a larger collection housed in SCUA, and the name of that collection is indicated in the object's description. Other digital objects in that collection are related to the item you are viewing in a variety of ways: either because of a shared subject, creator, or history of ownership. Once you have found an object of interest in Credo, you may wish to see other objects in the same collection. To do so, simply select the collection title found at the top of the item page and you will call up an entire list of every item in that collection. In the future we will be testing other ways to navigate the content of our large and varied collections within Credo.



DU BOIS ONLINE
"I BELIEVE IN LIBERTY FOR ALL MEN"

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011

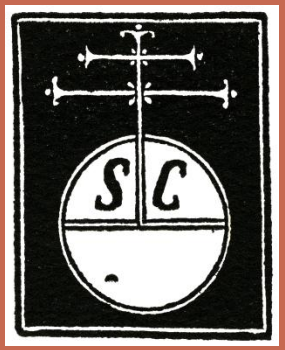
STANDARDS

The basic goal in designing Credo is to provide simple and intuitive access to digital collections within a repository architecture that is both robust and flexible and that adheres to open standards. The flexibility of the underlying Fedora Commons Repository software (maintained by the DuraSpace Organization) allows us to continue to develop Credo and to add additional features as we explore the best ways to deliver digital content to the public.

To provide richly detailed descriptions and finely-tuned discovery tools, we used the Metadata Object Description Schema (MODS) to describe all objects in Credo. SCUA is currently developing an extensive set of best practices guidelines detailing our implementation of MODS for the description of manuscripts at the item-level.

Although Credo is not a digital preservation system, we have made sure to use open, platform-independent standards and have maintained detailed information about our digital objects, which we have structured in metadata and outlined in external documentation. As the UMass Amherst Libraries develop a full digital preservation infrastructure, we are confident the digital collections in Credo will be available far into the future.

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011



RIGHTS

Reproduction and Use
Digital content made available in Credo is intended for personal research and educational purposes; commercial use or distribution of this content is not permitted without prior permission of the copyright holder.

The University of Massachusetts holds copyright to select content in Credo. Copyright to other content may be held by authors or their heirs, or may be in the public domain.

Requests to reproduce, distribute, or publish content in Credo should be sent to SCUA at credo@library.umass.edu, unless otherwise noted.

Takedown Policy
Individuals who wish to contest the use of specific works in Credo should contact SCUA. If you are the copyright holder of a work in Credo and wish to have that work removed, please submit a request for that work to be taken down. Your notice should include proof of copyright ownership and clear identification of the work (preferably with URL).

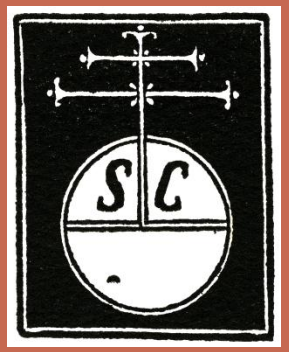
Resources on Copyright

United States Copyright Law

Fair Use

Copyright Term and the Public Domain in the United States

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011



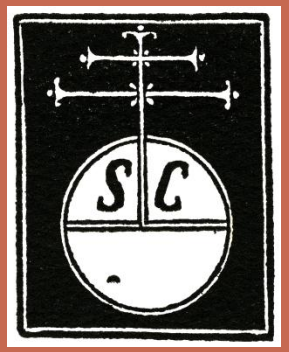
ACKNOWLEDGMENTS

Building Credo was a team effort, and SCUA was fortunate to have assembled an excellent team of archivists, programmers, and student assistants to contribute to this monumental project. Before any digital content can be loaded into a digital repository the skeleton that keeps this content must be constructed. The UMass Amherst Libraries Systems staff, M.J. Canavan, Aaron Addison, and Kelcy Shepherd, were instrumental in developing this structure and Josh Silver implemented Credo's design, created by Gravity Switch.

If the Systems Department built the body of Credo, then SCUA's Du Bois Digitization team supplied its heart. Aaron Rubinstein, project manager, led the team with creativity and innovation for the first two years laying the groundwork necessary for all our future success. Jeremy Smith, his successor, picked up the baton and kept the project on task and on schedule until we reached the finish line. Rob Cox, Department Head, and Danielle Kovacs, Curator of Collections, worked in concert with the project staff to make this shared vision a reality.

Providing access to digital collections is impossible to do without first creating accurate and detailed descriptions, or metadata, for those resources. Our early metadata specialists helped us to refine the standards and best practices we continue to use today: Abby Baines, Danielle Federa, Pamela Harnois, Brianna Haskins, Emily Toder, and Caroline White. Of course, there would be no digital content without high quality scans of the originals, and we employed a large team of committed undergraduate students for this task.

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011



PRESENTER BIOGRAPHIES

Jeremy Smith, M.S.,

Jeremy Smith is Digital Project Manager of the W. E. B. Du Bois Digitization Project at Special Collections and University Archives at the University of Massachusetts. He is a 2009 graduate of the Simmons College Graduate School of Library and Information Science. Before earning his MLS he was a producer/editor at the Media Education Foundation (MEF), a non-profit video production company specializing in educational videos. MEF videos explore contemporary social messages through the lens of the mainstream media.

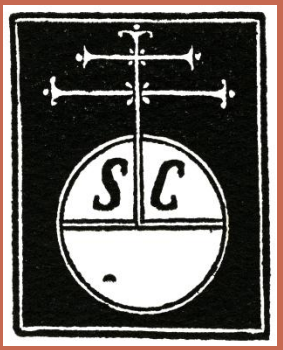
jlsmith@library.umass.edu

Abigail S. Baines, M.S.,

Abigail S. Baines has recently joined the Circulation and Reserves department at the W. E. B. Du Bois library as the E-Reserves Assistant managing online course materials for the University. Since May 2009 she has been a Metadata Specialist for the W. E. B. Du Bois digitization project in Special Collections and University Archives at the University of Massachusetts Amherst and a Library Assistant for the Goodwin Memorial Library in Hadley. She is a 2010 graduate of the Simmons College GSLIS program in Archives Management. Prior to Simmons, Abigail studied Communication and Psychology at UMass Amherst.

abaines@library.umass.edu

FROM W. E. B. to WEB:
Digitizing the Manuscript Collection of
W. E. B. Du Bois
Poster Session - ALA 2011



FROM W.E.B. TO WEB

Digitizing the Manuscript Collection of W.E.B. Du Bois

Special Collections & University Archives, UMass-Amherst

METADATA CREATION

3 Metadata records are created in the oXygen XML editor and encoded using the Metadata Object Description Schema (MODS). Local best practices for descriptive metadata have been developed for the item-level cataloging of manuscripts. Implemented standards include: AACR2, AAT, DACS, LCSH, plus local name and subject authorities.

INGEST

5 MODS records are wrapped in the Metadata Encoding and Transmission Standard (METS). There is a METS record for each digital object. The METS record also defines the digital objects by describing which image files make up a single object and stores technical metadata about the individual image files. All files are then ingested into a digital repository based on Fedora, the open source digital repository framework.

PROJECT FACTS

- The collection is being described at the item level
- There are approximately 100,000 items in the collection.
- The dates of the collection range from the 1870s to the 1960s.
- Portions of the collection were accessioned in the 1960s.
- As of the Fall of 2010, 22,000 metadata records have been created and 70,000 images scanned.
- FUNDING: comes exclusively from The *Andrew Mellon Foundation and the National Endowment for the Humanities*
- PERSONNEL:
  Project Director: 1 FTE
  Metadata Specialists: 2 FTE
  Scanners: 3.5 FTE (5-10 part-time positions)

RE-HOUSING

1 Original materials are re-foldered and re-boxed

SCANNING

2 Items are scanned as 300 dpi TIFF files and given a unique filename, e.g., mums312-0034-000-001

QUALITY CONTROL

4 Digital objects are checked for errors by both the Metadata Specialists and the Project Manager. The Project Manager uses Python scripts to analyze both the image scans and the metadata records.

USER ACCESS

6 A user interface, **Credo**, will provide access to the entire Du Bois collection, as well as all of SCUA's digital materials. The interface features item and collection level description, simple and advanced search, and faceted browsing.

PROJECT WORKFLOW

The "FROM W. E. B. TO WEB" Poster – Presented at ALA , New Orleans – 6/25/2011 - Abigail Baines & Jeremy Smith