White Paper Report

Report ID: 111912

Application Number: PW-51645-14

Project Director: Eric Pumroy (epumroy@brynmawr.edu)

Institution: Bryn Mawr College

Reporting Period: 5/1/2014-5/31/2015

Report Due: 8/31/2015

Date Submitted: 8/31/2015

HISTORY OF WOMEN'S EDUCATION OPEN ACCESS PORTAL PROJECT

WHITE PAPER

National Endowment for the Humanities
Division of Preservation and Access, Humanities Collections and Reference
Resources

Grant PW-51645-14

Project Director: Eric Pumroy Bryn Mawr College August 31, 2015

TABLE OF CONTENTS

1.	introduction p.3
2.	The Work of the Project
	A. Portal Design & Implementation p.5
	B. Metadata p.6
	C. Content p.8
	D. Usability Testing p.10
	E. Publicity p.11
3.	Advisory Board
4.	Ongoing Issues p.14
	A. Maintenance of the Site p.14
	B. Standards p.15
	C. Building Content & Use p.15
	D. Governance p.16
5.	Conclusion p.17
APPE	NDICES
Appendix A:	Project Description & Requirementsp.19
Appendix B:	Technical Development Documentation . p.21
Appendix C:	Metadata Standards p.24
Appendix D:	Report on Usability Testing p.26
	Project Participants

HISTORY OF WOMEN'S EDUCATION OPEN ACCESS PORTAL PROJECT

WHITE PAPER

1. INTRODUCTION

Digitization has become an essential component of the work of libraries and cultural institutions over the last decade, and as the process of creating digital libraries has matured, opportunities for connecting related but physically separate collections have become increasingly attractive. This kind of aggregation has been done on a large scale through comprehensive national collections, such as the Digital Public Library of America or the Internet Archive; or state or regional networks, such as the California Digital Library; or by vendors offering digital collections for sale. The purpose of the History of Women's Education Open Access Portal Project is to explore the potential for creating a collaborative, open-access, subject-specific digital site, one capable of expanding to include new collections and new institutions, and stimulating new research in the field through digital exhibits, social media, and links to new sites and publications. As a group of some of the oldest women's colleges in the country, we have found the need for this type of subject-specific site to be especially compelling and timely. Our institutions all have extensive holdings of letters, diaries, scrapbooks and photographs written or created by students, and together they constitute an invaluable, but underused, source for studying the lives of the first generations of American women who went to college and aspired to lives engaged in the public sphere. Gaining a greater understanding of these women's lives at college and beyond is essential for understanding the larger social, economic, political and cultural changes that took place in the United States over the last century, and will illuminate the ongoing debates about women's education and gender equality that continue both here and internationally.

While our primary goal in undertaking the project was to determine if we could find a way to connect collections in a vital and sustainable site, and over the long term, use that site to stimulate significant new work in women's history, we also saw a secondary goal of creating an open-source infrastructure and set of procedures that could be adapted by other institutions interested in developing their own subject-based digital collections. This White Paper, its accompanying standards documents, and the site architecture, design and accompanying documentation available through Barnard College's GitHub, are all readily available for other initiatives interested in pursuing this approach.

The portal project had its beginnings in 2012 when the director of The Albert M. Greenfield Digital Center for the History of Women's Education at Bryn Mawr College, Jennifer Redmond, began working with archivists from most of the other Seven Sisters institutions to create a digital exhibition on the topic of early entrance examinations at those colleges. Her contacts led to informal conversations about the similarities of our collections and the potential for expanding research opportunities if the collections could be brought together digitally. The conversations eventually led to a series of planning meetings, both online and in-person. As a result of those meetings, the institutions developed a planning grant proposal for the Foundations program of the National Endowment for the Humanities. The proposal was submitted in July 2013, and the award was announced at the end of March 2014.

During the course of the planning grant project, we focused our energies in two areas: 1) the development of an attractive, feature-rich website that would bring together the digital collections of the seven institutions; and 2) the development of standards and processes for managing our digital collections so that search results would be useful to scholars, students and other researchers, and so that the resources available through the site could continue to grow in a consistent manner past the life of the Foundations program grant.

This White Paper will outline the steps we took to design, build and evaluate the site and to create the standards and work processes to ensure that the site would be functional. As might be expected with planning grants, we came across a number of issues that we had not anticipated, and several that could not be resolved within the time and budget constraints of the project. The most important and difficult of these issues revolve around the questions of how a collaboratively-built site can be sustained over time, how it can be governed and managed, and how it can grow to reflect the interests of the researchers who will be its primary users. At this point we have done little more than raise these long-term questions, since our efforts during the project were focused on creating the initial site. Over the coming year, though, addressing the governance and sustainability issues will be the critical next steps. Fortunately, we have now been working together for more than two years, and over that time we have developed effective working relationships, strong respect for our colleagues at the other institutions, and a pride in seeing what we could accomplish through this first collaborative effort.

The project has been enormously successful. As the result of our work, we now have a dynamic and compelling website, *College Women: Documenting the History of Women in Higher Education* (www.collegewomen.org), that contains more than 300 digital objects from our institutions, and which has continued to grow since its launch in May. The initial reactions to our announcements and presentations have been overwhelmingly positive, with enthusiastic support coming in from across the country from both scholars of women's history and professionals active in digital scholarship. We recognize that there is still much work to be done

to build the collections and create the curated content that will make the site truly useful for research and education, but the project has laid the essential foundation for making those developments possible.

The production of this report has been a team effort based at Bryn Mawr, but with contributions from many other members of the collaborative. Particular thanks go to Rachel Appel for her work on the Metadata section, Evan McGonagill for the Usability Testing section, Monica Mercado for much of the Content and Publicity sections and some of the introduction and conclusion, and all three of them for reviewing and editing the entire document. In addition, commentators Joanna DiPasquale, Sarah Goldstein, Helen Horowitz, Elisa Lanzi, Laura Streett, and Martha Tenney all made significant contributions to improving the text of the report.

2. THE WORK OF THE PROJECT

Most of the work on the project was done by subcommittees working together by email, telephone and Google Hangouts. In addition, there were four project-wide meetings at which key decisions on the site were made. The first of these meetings, in May 2014, was done virtually using Google Hangouts. The other three were at Barnard in July, Vassar in December, and Smith in March. Members of the Advisory board also attended the December and March meetings, most of them virtually.

The development of the *College Women* site focused on the following areas: A) Portal Design and Implementation; B) Metadata; C) Site Testing; D) Content of the site; E) Publicity.

A. Portal Design and Implementation

When we wrote the proposal in the summer of 2013, we anticipated that we would need to evaluate a number of content management systems for displaying text and images as part of the initial phase of the project. Over the course of the following year, though, we watched with interest the development of the Digital Public Library of America, and on a much smaller scale, the Brooklyn Visual Heritage Project, and the successes of those initiatives led us to rethink the structure of the site. The Brooklyn Visual Heritage Project in particular looked like a promising model since, like us, it was an attempt to bring together digital collections from multiple institutions, each with their own pre-existing digital collection repositories and metadata systems. Like us, too, they recognized that the cost of building and maintaining a new digital repository capable of storing and managing archival-quality images would be not only prohibitive, but also unnecessary since all of the institutions had their own systems for managing digital collections. Instead, the approach they took was to build a site that harvests metadata and small images from the participating institutions' digital repositories, creates a structure for searching the metadata and displaying the results, and then links the user back to the higher-

quality images on the institutions' websites. They built the site in Drupal, which was also attractive to us since many of our institutions are already using Drupal for our websites, and so the language and structure was familiar to us. After a number of conversations with staff who had played important roles in the development of the Brooklyn Visual Heritage Project, we decided that this was the appropriate model to follow.

To serve as the developer on the project we selected Interactive Mechanics, a Philadelphia firm that Bryn Mawr has worked with successfully on several other projects, and which also has extensive experience working in Drupal. During the early part of the summer the group developed the functional requirements for the system (see Appendix A), and these were used by Interactive Mechanics to prepare preliminary designs. At our late July meeting at Barnard College, we approved these designs, and discussed a number of the technical issues that would still need to be resolved, particularly those around harvesting data from our individual sites.

Over the course of the fall, Interactive Mechanics developed the website and did test harvesting of metadata and thumbnail images from Bryn Mawr's digital collections. At the group's December meeting at Vassar, the more advanced version of the site was discussed and approved, and a process was put in place to load more images and data and begin testing the site in early February. Further changes to the site were made in the spring as a result of the testing, particularly to simplify the search process and establish direct links from the small images in the portal to the larger images in the institutions' digital repositories.

One of the critical factors to the success of the project was our good working relationship with the site's developer, Interactive Mechanics, and its director, Michael Tedeschi. Rather than simply carrying out our instructions, Mr. Tedeschi was an active participant in the discussions about how best to achieve the functionality we were looking for. His presentations to the group and follow-up discussions served as the starting points for each of our project meetings, and he participated in numerous conservations throughout the year as we worked on improvements to the site's functionality. Having a developer who was an active partner in the planning process made for a much better site.

Since one of our project goals is the creation of a model for subject-specific digital collections, the site was designed to be open-source and shareable so that others can make use of our work. To make site architecture easily accessible, we have made our Drupal theme and associated files available at Barnard College's GitHub repository:

https://github.com/BarnardArchives/brynmawr-collegewomen-drupal. In addition, the technical development documentation is attached to this report as Appendix B.

B. Metadata

Finding ways of standardizing the metadata was one of the first critical challenges we faced in the project. All of the institutions had been building digital collections for many years, and each had its own digital repositories and standards for cataloging digital collections. In order to find a way of bringing this metadata into a coherent structure, we created a "data management" team at the beginning of the project which focused on developing a metadata schema and data dictionary with a controlled vocabulary.

During our July 2014 meeting, we agreed upon a metadata structure based on Dublin Core, in addition to a few local fields such as "Institution" and "Subject-Local" that would enable users to perform both faceted and keyword searches and provide contextual information. (see Appendix C). Qualified Dublin Core was selected because of its flexibility for mapping and common use across repositories. Over many discussions, the data management team worked collaboratively and created a data dictionary which included the metadata fields, their definitions, whether the field was repeatable, and any controlled vocabulary. The different institutions all use adapted versions of standard metadata schemas, such as MODS, Dublin Core and VRA Core, and the data management team mapped them individually to the standardized site schema. We decided to create a list of "themes" or topical subjects to better connect our materials curatorially and enable scholars and students to pull together related results easily. The data management team created a list of themes such as traditions, friendship, and student movements. Many of us were also able to utilize the APIs (application programming interfaces) at our individual institutions and were able to directly import reference URLs that link back to the item on its local repository and pull images and thumbnails from the original location without uploading any actual files to the website.

We ran into several issues. Our initial goal was to use the Drupal module Feeds OAI-PMH Fetcher and Harvester to create an automated process to harvest our metadata from our repository APIs, transform it to fit the portal standard, and ingest it to the site. However, because of our closed-source local digital asset management systems, there was very little flexibility for conversion. For those institutions with proprietary systems, or systems without an API, the metadata would only export in the structure it was already in without converting. This placed a burden on the colleges that worked with these systems; without programming/technical support on their end, they could not make their data interoperate very easily. In addition, Feeds OAI-PMH Fetcher and Harvester would only "fetch" metadata that fit within traditional Dublin Core, rather than qualified Dublin Core. For example, the OAI Feeds Module only includes the main 15 Dublin Core fields but all seven institutions use at least one or two qualified versions of those fields. In one case the "Format" field was defined as "medium, genre, or type," a qualified field which the module would not ingest at all because it did not recognize it as an exact match. There was no way to modify the module to accept different, mapped fields. In order to include metadata in a consistent way, we made the decision to scale back our item ingest, and upload our

metadata manually in TSV (tab-separated value) files, essentially an Excel spreadsheet, which we then batch imported into Drupal. Being able to automatically harvest metadata is the most sustainable way to keep the records in sync, and the next phase of the portal project will look at building an entirely new OAI-PMH metadata harvester that is flexible and can convert metadata into any schema you describe into the module.

Normalizing metadata was another challenge. Each institution's local standards were not an exact fit, despite mapping it to the closest relevant field in the portal's schema. Institutions have local practices and we drew subject headings from LCSH (Library of Congress Subject Headings), TGM (Thesaurus for Graphic Materials), and locally developed ones into one "Subjects" field. While this aggregation provided a level of flexibility to the "ingest" phase of the project, it also meant that aggregation of terms was more difficult once we received all of the data since it was difficult to provide connections between similar but not exact terms. For example, institutions used different standards in defining photographs for the field "Format," with some using just "Photograph," and others "Black-and-white photograph," or "Photographic print." Some institutions used pre-coordinated local subject headings, with the name of the college followed by the subject term. An example of this is "Bryn Mawr College--Students" which can only link to other Bryn Mawr College records. Our decision to include larger "theme" subjects proved useful at this point, allowing local practice to remain in place while providing a pathfinder to end users by mapping on terms that they were likely to be interested in and that many items in the aggregated collection had in common. After ingest, we did some metadata cleanup, but ran into the question of how much manual cleanup labor we should be doing for a site that is meant to be a portal and not a repository. For the next phase, we have to consider the balance between optimizing search and discovery for a particular record set and optimizing it for aggregation. There may be a way to automate cleanup through the new module, but we hope to streamline this process through revisiting the metadata schema and having further discussion on the use of subject terminology.

C. Content

At the July and December 2014 meetings of the seven institutions, initial discussions were held about the date ranges and types of materials appropriate to harvest for the portal, the portal's official title and explanatory texts, and supplementary materials to offer as part of the portal's links and resources. A "content group" was set up during our December 2014 meeting at Vassar College in order to lead the decision-making process on these issues, and to write the explanatory text for the site.

In order to support the work of harvesting materials for the portal, the institutions first made decisions about what materials to harvest for the initial, beta site launch. We decided to encourage the partner institutions to focus on the years before World War II as a way of building

a preliminary collection that had some subject coherence across our colleges. Although most institutions had tremendous wealth of photographs, we also made a concerted effort to vary the materials harvested, allowing us to experiment with compound objects--collections of letters and scrapbooks in particular.

As discussed in the Metadata section, there were complications with subject headings because of the variations in style and standards between institutions. Because there were headings from LCSH, TGM, as well as local subject heading indices, there was difficulty searching for similar or related items across the collections. Although there are still issues to be addressed, one solution was the addition of themes. In essence the set of themes is a simplified local subject heading index - local to the *College Women* portal. The list of themes was developed from the topical strengths of the existing portal material, as well as from topics we hope to document further. Future discussion on improved searchability will involve crafting a firm set of definitions for the themes, so they can be more easily applied; considering what LCSH phrases we might favor over others; and whether local subject headings should be harvested for the portal.

After discussion with the partner institutions and the project's advisory board, we settled on the website name *College Women: Documenting the History of Women in Higher Education*. There was agreement that the site's name should not reference the "Seven Sisters," since we would like to explore future collaborations with a wider range of institutions charged with educating women. The language of the site does not restrict its future use to just the schools known as the "Seven Sisters," but positions the portal as a critical project for the study of the history of women college students. We also decided that the site title, tagline, and introductory texts needed to be clear that the site leads to materials relating to student experiences, and not administrative or curricular histories. Site texts, including title and tagline, expanded description about the site, and a blog, were drafted in time for the February 2015 testing process.

In April 2014, after usability testing, the content group proposed a number of items necessary for the site to go live in May 2015. These included a site URL (collegewomen.org), shortened introductory text for the homepage, images to illustrate the rotating slideshow of themes on the homepage (Arts, Theater, and Music; Academics; Traditions; and Sports), and an "About" page that would summarize the project's goals, history, and contributors. At this step in the process, the group decided that a proposed "Help Guide" for users would not be necessary until the site grew considerably.

Testing feedback also made it clear that more curated content would be useful to the site's users. The content group spent their remaining time considering what other kinds of resources we might create and where it should be housed on the site. For the May 2015 launch, we started by adding site information to a blog, along with sample posts about college history. We also plan to work with our partner institutions to share the history of each school, and the relationship of the

"Seven Sisters" to each other (a question that came up in site testing), in dedicated blog posts integrated into the site. The content group also created a simple Links & Resources page in order to direct site users to digital collections of each of the seven partner institutions, other major digital collections and exhibits on women's education from other U.S. institutions, and a bibliography of further reading.

A future goal would be a custom installation of Omeka, the digital publishing platform, making curated exhibits possible as part of the site. Moving forward, we hope to think about digital exhibitions, a more robust bibliography and site links, and scholarly essays. A Bryn Mawr College history seminar in the spring of 2015 used the pre-launch site for the students' primary source research into women's higher education, an experience that suggests possibilities for expanding content aimed at educators who might use the portal in their classroom.

D. Usability Testing

A "usability testing group" was also set up during our December 2014 meeting, consisting of four team members from Bryn Mawr, Barnard, Mount Holyoke, and the Radcliffe Institute. The group was charged with evaluating the forthcoming prototype and delivering feedback to the rest of the group as well as to our developer, Interactive Mechanics, for revisions to the content and functionality of the site. Between December 2014 and the launch of the site in May 2015, we performed our charge in five phases: research, test design, testing, results analysis, and implementation of recommendations.

We began the process by speaking with a professional usability consultant and other colleagues experienced in this area to help define our testing approach, eventually settling upon a method in which we would hold sessions with individual participants using a uniform protocol and script to ensure consistent data collection. The user would be invited to explore the site and asked to perform specific tasks that were essential to site functionality, and encouraged to offer verbal feedback on their experience as they moved through the interface. The representatives on the testing team conducted eighteen sessions across the four institutions with participants who were staff, faculty, students, and others at and beyond the four schools.

Broadly, our results showed that users were compelled by the site concept and design, but struggled with content discovery. The variety of methods that the site offered for entering and exploring the collections (keyword search, advanced search, browse, an "Explore the Collections" slideshow on the front page) tended to confuse users, who wanted to see a unified and visually clear entry point. We also found that the visual layout and textual content of the homepage did not effectively communicate the mission or functionality of the site. In addition to several small tweaks to the interface, we communicated these main points of feedback to Interactive Mechanics and worked with them through several iterations to settle on designs that

were more intuitive, more efficient, and did a better job exposing and interlinking the collections. User feedback also revealed several instances in which the textual content that had been written for *College Women* by the content group could be adjusted to be clearer or more appropriate in length. These changes were implemented in time for the May launch of the site at the Women's History in the Digital World Conference at Bryn Mawr College.

There were several desired features that we brought to the attention of the group and the development team that we resolved not to work on during this phase of the project, though they may be candidates for future development iterations. Certain search features would require more negotiations among the metadata team in order to clarify standards between different institutions, such as the ability to narrow results by item format. We were also limited in some instances by our platform, such as the site's inability to specify whether to use exact phrases or all words included in a search term, a feature that is not included in our current Drupal search module. The commonly expressed wish for more robust contextualization of individual collection items is a goal for future versions of the site, but will require further conversations within the project team about what types of context would add the most value to the site, and a significant amount of staff time will need to be dedicated to content creation. While these features were outside the scope of the prototype stage, several of them may be addressed in the upcoming phases of development.

A full summary of the testing process can be found in Appendix D.

E. Publicity

College Women: Documenting the History of Women in Higher Education was informally launched at the May 2015 conference Women's History in the Digital World, sponsored by The Albert M. Greenfield Digital Center for the History of Women's Education. The content group drafted a press release that was shared with each of the partner institutions shortly thereafter, and each of the seven schools organized publicity aimed at its own community and alumnae. As an example of the institutional media, the Bryn Mawr College website had the College Women site as a featured item for most of the summer (http://news.brynmawr.edu/2015/07/09/seven-sisters-partners-launch-new-archives-project/), and there will be a story on the site in the fall issue of its alumnae magazine.

Examples of the other institutions' publicity include:

Barnard: http://library.barnard.edu/news/seven-sisters-partners-launch-new-archives-project

Mt. Holyoke: https://www.mtholyoke.edu/media/seven-sisters-site-chronicles-women-higher-ed

Schlesinger Library, Radcliffe Institute, Harvard:

http://www.radcliffe.harvard.edu/schlesinger-library/blog/collaboration-among-seven-siblings

Smith College, Friends of the Library Newsletter:

http://us3.campaign-archive1.com/?u=249657c81b98fb61ac3921497&id=85e3167b18

Vassar College: http://library.vassar.edu/announcements/announcements/2015-2016/150723-8evenSisters.html

The project was also announced via social media, electronic lists, and newsletters, and these announcements have generated a large number of comments on social media. To review them, see Storify: https://storify.com/GreenfieldHWE/launching-college-women.

Announcements were sent to, or appeared in, the following:

Electronic Lists

H-Education

H-Women

H-DigHistory

Philly DH

MARCH (Mid-Atlantic Regional Center for the Humanities)

PACSCL-L (Philadelphia Area Consortium for Special Collections Libraries)

ExLibris

SAA WCRT (Society of American Archivists – Women's Collections Roundtable)

SAA College & University RT

NCPH (National Council on Public History)

NCWHS (National Collaborative for Women's History Sites)

CCWH (Coordinating Council for Women in History)

Religion in American History blog

Technically Philly: http://technical.ly/philly/2015/06/22/seven-sisters-digital-archive-bryn-mawr/

Print publications

Archival Outlook (Society of American Archivists)

ACRL News

American Libraries

Because the site launch took place during the summer, further outreach will be necessary to potential researchers--including scholars, students, educators, and women's college alumnae/i.

Project team members are also committed to sharing the site at a number of history, archives, and digital humanities conferences this year. This year's talks include:

- Nanci Young and Elisa Lanzi (Smith College), "The History of Women's Education Open Access Portal Project," ARLIS/NA New England Chapter Meeting, October 2014
- Rachel Appel (Bryn Mawr College), "College Women: Documenting the History of Women in Higher Education," presented at the Women's History in the Digital World Conference digital showcase, Bryn Mawr College, May 2015.
- Leslie Fields (Mount Holyoke College), "Digital Humanities and Special Collections: New Tools, Challenges, and Opportunities," Rare Books and Manuscripts Preconference, Cleveland, OH, June 2015.
- Monica Mercado (Bryn Mawr College) and Michael Tedeschi (Interactive Mechanics), "Documenting the History of Women in Higher Education," presented at Keystone Digital Humanities, Penn Libraries, July 2015.
- Monica Mercado, "College Women," scheduled for the Upstate New York Women's History Organization conference, Seneca Falls, New York, September 2015.
- Rachel Appel, Joanna DiPasquale (Vassar College), and Monica Mercado, "College Women: A Collaborative Cross-Institutional Archives Portal," scheduled for the Digital Library Federation Forum, Vancouver, British Columbia, October 2015.

The initial reactions to our announcements and presentations have been overwhelmingly positive, with enthusiastic support coming in from across the country, as demonstrated by the social media comments collected on Storify. We have already heard from several faculty members who plan to use the site in their courses in the coming year. We also had a response from a Black alumnae group from one of our colleges asking how they could contribute to the site, a promising indication that *College Women* has the potential to promote the preservation of more student writings, particularly from the second half of the twentieth century.

3. ROLE OF THE ADVISORY BOARD

The members of the project's advisory board played a critical role at the two meetings they attended virtually—at Vassar in December and Smith in March. When so much of the project's work was focused on technical issues around the portal development and metadata, they provided strong voices to remind us that the success of the project ultimately depended upon its value to scholars and students. They were particularly helpful in sharpening the definition of the portal to the history of the student experience, rather than larger college administrative history,

and in keeping the definition of the site open enough to allow for the eventual expansion of the site to women's colleges beyond the Seven Sisters.

4. ONGOING ISSUES

The launch of the College Women site was the result of a year-long process of planning, discussions, testing, and document-writing that involved more than twenty people at our seven institutions and members of the Advisory Board. Our success in building a functional site demonstrated the ability of the group to work together effectively toward an immediate goal, but it now leaves us with the challenge of devising mechanisms for maintaining and governing this site over the long term. Complicating this challenge is the pilot nature of the site. It is difficult to make a case for devoting significant time and energy to building governance structures and committing institutional resources to a site that hasn't yet proven its value, and still requires considerable development and enrichment before anyone is likely to find it useful. Rather than focus on governance immediately, we made the strategic decision in the spring that the best way of moving the project forward was to concentrate instead on grant proposals that would bring in the money needed to build the site's content and improve its functioning. Now that proposals have been submitted to the "Digitizing Hidden Special Collections and Archives" program of the Council of Library and Information Resources (CLIR) and to the Humanities Collections and Reference Resources program of the National Endowment for the Humanities, we have returned to the governance, maintenance and development issues, and will be addressing them over the coming year, as well as pursuing other funding options for digitization.

The principal long-term issues are: 1) maintaining and enhancing the site; 2) assuring adherence to standards; 3) building audience for the site through enriching the content and promoting its use; 4) finding an effective governance structure.

Maintenance of the site. Interactive Mechanics developed this Drupal-based site on its own servers, and turned it over to Bryn Mawr at the completion of the NEH planning project. Because Interactive Mechanics is still providing support for the site and staff members at each institution need access for uploading new records, the site could not be housed within Bryn Mawr's secure network. Instead, the site was loaded onto a cloud service, Amazon Web Services, where it also has a basic level of support from DLT Solutions. At this pilot stage of the site the costs for hosting and support are modest, and the initial fees were paid with funds from the NEH planning grant. As the site's collections grow and its use expands, though, these support and storage costs will also increase. Most of the institutions indicated a willingness to share in the costs of maintaining the site beyond the grant period, but that willingness was also

based upon the test site's low costs. As the site grows and the costs rise, the site will have to prove its value to the participants in order to justify increasing institutional investments. At some point, too, it may become cost-effective for one of the institutions to take over the hosting and maintenance of the site, supported by payments from the other participants.

Standards. As discussed in the Metadata section of the report, one of the priorities of the coming year will be working on normalizing the ways in which we describe collections and use subject terminology so that users' searches produce meaningful results. Further discussion is also needed on the handling of compound objects – scrapbooks, diaries and collections of letters. Exposing the rich content of these complex documents depends upon extensive, item-by-item description and links that lead users directly to the relevant sections of the documents. Several institutions have also been creating searchable transcripts of documents as a way of making their content even more accessible. We included a mechanism for displaying transcripts in the portal (see, for example, Vassar's Sarah Bates Letter, 1868), although this display works best for relatively short documents, such as letters, and much less well for lengthy pieces, such as diaries. At this nascent stage in the portal's development we still haven't had enough experience with the various approaches for exposing the detailed content of lengthy complex documents to know which ones will work best. Further experimentation and discussion will be an important part of our work in the coming year.

Building Content and Promoting Use. For the purposes of the planning grant we have defined success as meeting the proposal's goal of launching a functional site that draws together digital collections from all seven institutions. In the long run, though, the success of our effort will be measured not by the functionality of the site, but by how the site is used by scholars, teachers and students, and whether it is able to stimulate new thinking and teaching on the history of women in America and the world. In its pilot state, the *College Women* site only shows its potential. To realize that potential, we will need to work on building the content of the site with both extensive digital collections and rich curated content that we have submitted several grant proposals that, if successful, will enable us to digitize large parts of our holdings in just a few years. If those proposals are unsuccessful, we will still undertake digitization of our collections using our own resources, albeit at a slower pace. Nonetheless, we think that with judicious choices of collections to digitize and a systematic approach to the work, the site should have enough content to be broadly useful within a year. The other essential component of the site needing development is curated content that suggests avenues for research and ways of using the collections for teaching and student projects. The principal avenue for presenting this curated content is through the site's blog. Several of the participants already do blog posts for their own institutional sites, so as a preliminary step we will be adding some of these posts to the College Women site. We will also plan to add posts as new collections are digitized, and we will be inviting short essays and teaching suggestions from scholars who have been following the project.

Most of the publicity for the site came out during the summer, at a time when many of the potential users are away. Accordingly, we will be doing additional publicity in the fall, particularly aimed at sites focused on women's history and the history of education. Over the longer term, we will need to build a structure for engagement with scholars. The project had an advisory board of scholars during the planning grant, and this is a promising basis for what might become a larger group of advisors who can help ensure that the site grows in a way that is reflective of the needs of scholars, teachers and students.

Governance. The project governance during the planning grant was very much an ad hoc approach to accomplishing the goals of the project while engaging the appropriate staff members from all seven institutions. The principal structure for the project was provided by the grant itself, which set firm deadlines and assigned the responsibility for seeing that they were met to the official grantee, Bryn Mawr College. Mindful of that responsibility, the project director and other staff at Bryn Mawr took the lead in setting deadlines and goals, working with the site developer, organizing meetings and setting agendas. While this leadership role was appropriate for a short-term project with specific goals that needed to be accomplished, we were all aware that it is not a suitable governance structure for a long-term collaborative initiative. Critical to the success of this collaborative project was the active engagement of staff members from all seven institutions. Because the project required expertise in both digital technologies and the collections of student writings, each of the institutions involved both their archivists and their digital collections specialists. In all, more than twenty people from the seven institutions were involved in the project's discussions and planning. The range of voices and perspectives was essential to the successful development of the site, and also played an important role in building the sense of a shared mission and commitment to the idea of a common site on women's college experiences.

Even though this loose structure worked for a short-term project, the *College Women* site's long-term success will require a more formal governing approach that reflects the consortial nature of this effort. Decisions about the ongoing structure were intentionally deferred at our March 2015 meeting at Smith because of the pressing decisions that needed to be made about the final enhancements to the site and the quickly approaching deadlines for grant programs that could fund the large-scale digitization that we need to undertake in order to build the content on the site. The writing of the new grant proposals has provided the opportunity to outline a tentative structure for governing those potential projects, and if one of the proposals is successful, the fleshing out of ongoing structures will be helped along by the regular meetings and a project manager. The main outlines of the structure seem clear, though. There will need to be an ongoing coordinating group of manageable size so that the group can meet virtually on a regular basis. There will also need to be a set of committees charged with working on specific aspects of the site, particularly metadata, functionality, and the site's content. Those groups will have

reporting lines to the coordinating committee, perhaps with committee chairs sitting on the coordinating group. In addition, we will want to maintain an advisory committee of scholars to ensure that the site develops in a manner that makes it useful for research and teaching. Finally, we will need to have a memorandum of understanding that outlines how decisions are made.

5. CONCLUSION

The idea for this project began with the concern that our seven institutions' extensive collections of college women's letters, diaries and scrapbooks were a large, important, but underused set of documentation on the history of women in the United States from the mid-nineteenth century to the late twentieth century, despite the rich stories they tell. Women's history--and particularly women's education history--has been underrepresented in the turn to digital humanities in spite of the enormous changes that have taken place in American society over the last century as a result of women's education and the successful struggle for women's rights. By working together, we aimed to address two major issues: providing free, open access to digitized materials related to the history of women's education, and piloting a model for multi-institution collaboration. We believe we have succeeded on both counts, although we have much more work to do to achieve the site's potential as a critical resource for not only women's history, but also the history of American higher education and American society in the twentieth century.

College Women lays the foundation for new teaching and research in the history of women's higher education in the United States, and for new critical scholarly analysis in the interaction of women's education and larger changes in American society. As our first forays into publicizing the project have shown us, the portal does more than simply offer a new way of accessing archival materials; by putting the collections of seven women's institutions in conversation with each other through thematic searching, we have the opportunity to demonstrate why the education of women was, and continues to be a central issue in women's history and American social history. Further development will allow us to include current students, as well as women's college alumnae as creators and curators of this cultural heritage, and to think about how this project can serve as a model for further subject-specific collaborative digital collection projects.

By digitizing and aggregating the collections selected by the seven partner institutions, this project will contribute hundreds of thousands of women's words to the corpus of digital archives presently available. Until recently, women's educational history, and indeed, women's history in general has been underrepresented in the digital revolution in humanities scholarship, and much of what is available online has been held behind the paywalls of proprietary collections. This project will grant access to a new body of rich source material for those interested in studying the national landscape through women's lives and educational experiences in the late nineteenth and

early twentieth centuries, and in so doing reveal stories and voices that are often absent from the cultural record. As these collections grow, and are discovered, they make the case for the importance of women's education, both in the past and today.

APPENDIX A

The History of Women's Education Open Access Portal Project Summary Description and Requirements July 2, 2014

The History of Women's Education Open Access Portal Project will create an online public database for finding digitized historical documents written by students at the colleges once known as the Seven Sisters (Barnard, Bryn Mawr , Mount Holyoke, Smith, Vassar, Wellesley, and Radcliffe – now part of Harvard).

The database will consist of metadata from the 7 institutions, including links to digital objects available at each institution's website; and thumbnail images from each digital document. The database will include an automated system for harvesting newly-created metadata from each institution.

The database will have a public interface which will include mechanisms for

- 1) searching the metadata and displaying the results,
- 2) explaining the scope and purpose of the project
- 3) highlighting project news

The project team has agreed on the following requirements for the project. The requirements are still under discussion and additional ones may still be developed, but these cover the core of the project:

Specific Site Functionalities

Searching (full-text)

Filters/limits

Results per page

Sortable search results

Browsing

Search by broad categories (place, time period, subject, names?)

Syndication for end-users/RSS Feed

Links to all of the institutions/contact info/contact web form

Creative commons info/materials use policy for each institution

Google analytics or other usage reporting

Twitter feed (we do not have to have an account, but perhaps a hashtag)

Gallery viewing

Social media share feature integration including email, Facebook, Twitter, Pinterest, etc.

Slideshow or several item display changeable interface

News/events section

Links to original records/reference links

External links use target = "_blank"

Similar items

Exportation of data/citations from galleries

Terms of service for users

Administrative functional requirements:

Admin login

Configurability to change: searching, themes, edit records

Minimal maintenance Run scripts for harvesting

User management; ability to approve, delete

Comment approval

Non-functional

Section 508 compliant

Usability

Sustainable

Audit and control

Documentation

Disaster recovery/backup

Extensibility

Scalable

APPENDIX B

TECHNICAL DEVELOPMENT DOCUMENTATION

The project is realized in a website designed by <u>Interactive Mechanics</u> using Drupal 7.3.2 using the following modules:

- <u>Chaos Tools (CTools) Suite</u> provides a suite of tools for developers to manage common Drupal features
- <u>Date, Date API</u> allows specialized date verification to ensure best practices
- <u>Features</u> allows developers to build components in a repeatable way (similar to macros for Microsoft Excel) by capturing common functions and building on them
- <u>Feeds, Feeds Tamper</u> enables harvesting and specialized parsing of our institutions' data
- Field Collections organizes specialized data types into groups or fields
- Prev/Next adds information to each guery about the sequencing of pages
- <u>Taxonomy Manager</u> provides a more useful way to manage a hierarchy of values
- CKEditor a common WYSIWYG editor for Drupal
- <u>Views, Views UI</u> an incredibly powerful module that provides access to structured data through interface-instructed database queries, allowing us to build lists, browsable items, and more.

Theme Specs

Built with Sass, iQuery

Third-party:

- Twitter Bootstrap
- FlexSlider
- o PDFJS

Drupal theme and associated files for the project can be found at Barnard's GitHub repository

FUNCTIONALITY

Homepage Functions:

- Slideshow featuring themes
- Blog
- Google analytics
- Contacts page
- Institution partners
- Funders

Search Functions:

- Large central search bar
- Keyword search
- Full-text search
- Advanced search
- Boolean search
- Autocomplete search terms

- Field restricted keyword searching
- Filters/limits through facets
- Browsing by theme

Results Functions:

- Results per page
- Sortable search results
- Page through results
- List view and gallery view

Item pages:

- Links to all of the institutions for item reference
- Social media share feature integration including email, Facebook, Twitter, Pinterest, etc.
- Links to original records
- Generated images through reference URL
- Recommended citation
- Linked subject headings
- Similar items

Non-functional:

- Section 508 compliant
- Usability
- Sustainable
- Audit and control
- Documentation
- Disaster recovery/backup
- Extensibility
- Scalable

NAVIGATION AND DESIGN

Homepage

- Search Bar: the homepage features a large central Search bar that performs a full-text search of site contents, and a link to an Advanced Search pop-up window.
- Themes Carousel: directly underneath the search is a slider featuring [six?] rotating images representing major themes found in the collections. Each item in the collections is tagged with relevant themes, and clicking on the images in the carousel will lead to a filtered search that displays a list of images found in that topical category.
- About the Collection: beneath the carousel is a brief statement about the scope and purpose of the site.
- Featured Posts from the Blog: an area beneath the site description displays dates, titles, and excerpts from the three most recent posts added to the blog.
- Contributor Logos: logos from all participating colleges link outwards to those institutions' websites.
- Footer Documents: the page footer features PDFs of our Terms of Use, a Help Guide for using the site, Technical Docs that explain the functional requirements and development process of building the site, and our License Policy for all site content.

Advanced Search

 The Advanced Search feature, which can be accessed by the link below the homepage search or the icon in the upper right hand corner of the navigation menu, allows a user to narrow by multiple facets to craft a more specific search. A Keyword Search option performs a full-text search with the same functionality as the homepage search bar. There are also options to narrow by date, subject, item format, and institution.

Browse

The Browse page defaults to a complete list of items on the site. A menu of major topics
found in the collections allows the user to explore by theme, and a link to the Advanced
Search pop-up window can narrow the listed results by various facets. The results can
be toggled between List view and a gallery-like Grid view.

Record Pages

Search or browse results lead to a list of items that, when clicked, bring the user to a
page for the individual record. Each record page displays an image of the item,
metadata, copyright information and citation, social media sharing options, a link to the
item's home repository, contact information for the originating institution, and a list of
similar items in the collection.

About

• The About page offers context for the website and collections, including blurbs on the history of higher education for women, the history of the project, a list of contributors, and contact information for all institutions regarding individual collection items or general questions about College Women.

Links & Resources

 The Links & Resources page supports further inquiry by providing links to the digital archives of all contributing institutions, as well as a bibliography of suggested further reading on the history of women's higher education in the United States.

Blog

 The Blog is a hub for news and updates on the project and original content written by staff at the contributing institutions featuring items from and history of the collections.

APPENDIX C Metadata Standards

A set of common metadata elements and protocols were developed by the seven partner institutions for the College Women project (www.collegewomen.org), as follows.

Field	Definition	Controlled	Repeatable
Title [required]	The name given to the resource; short and descriptive.	None	no, unique
Creator	An entity primarily responsible for making the content of the resource.	If a person, Last Name, First Name (role; if applicable)	yes
Description	An account of the content of the resource; longer and more detailed description than Title.	none	no, unique
Location	Geographic coverage.	none (recommended best practice: TGN, http://www.getty.edu/research/tools/voc abulary/tgn/index.html)	yes
Display Date	Date created, circa date, date span	circa spelled out	no, unique
Start Date	For backend search parameters	YYYY for Start Date/End Date	
End Date	For backend search parameters	YYYY for Start Date/End Date	
Subject Local	Local to Institution (includes all from local repository) The topic of the content of the resource.	none; use TGN when applicable for place names: http://www.getty.edu/research/tools/voc abularies/tgn/index.html	yes
Themes	Prominent topics from the collections of potential research interest: Academics; Arts, Theater, and Music; Buildings & Grounds; Celebrations; Class Reunions; Commencement; Dormitories; Dress; Events; Faculty; Friendship; Firsts; Political and social activism; Race; Recreation; Religion; Societies; Sports; Student activities; Student movements; and Traditions	yes, from list	yes
Contributor	An entity responsible for making contributions to the content of the resource.	none	yes
Format		http://www.loc.gov/standards/valuelist/ marcgt.html	yes
Type [required]	The nature or genre of the content of the resource.	http://dublincore.org/documents/dcmi- type-vocabulary/	yes
Physical description	,	none	yes
Institution [required]	you would like it to appear.	none	no, unique
Rights	Information about rights held in and	none	yes

[required]	over the resource.		
Contact	Boilerplate statement directing the user to the Contact page.	none	no, unique
Reference URL [required]	Callback URL to local repository	none	no, unique
Image File Path [required: unless audio file]	Callback URL to larger image file	none	no, unique
Thumb File Path [required: unless audio file]	Callback URL to thumbnail	none	no, unique

APPENDIX D REPORT ON USABILITY TESTING AUGUST 2015

A "usability testing group" was also set up during our December 2014 meeting, consisting of four team members from Bryn Mawr, Barnard, Mount Holyoke, and Radcliffe. The group was charged with evaluating the forthcoming prototype and delivering feedback to the rest of the group as well as to our developer, Interactive Mechanics, for revisions to the content and functionality of the site. Between December 2014 and the launch of the site in May 2015, we performed our charge in five phases: research, usability testing design, testing, results analysis, and implementation of recommendations.

Research: because none of our testing group members had formal experience in usability testing, we began the project with a research phase. This included a literature review in which we surveyed a wide range of approaches to usability testing. Though our project was influenced by many of the sources discovered in this phase, we used texts by Jakob Nielsen and Steve Krug as key references while designing the tests. We also had a series of conversations with usability consultant Kelly Mueller, who coached us on some of the best practices and guidelines for designing and implementing a testing protocol. After the research phase and the drafting of the testing script, we sent the Usability Test Plan document to the rest of the team and incorporated their suggestions before proceeding. Throughout the process, we submitted planning documents to and solicited feedback from the larger project team.

During this phase we also discussed accessibility and researched methods of evaluating the site for Section 508 compliance and enhancing ease of use for the visually impaired. We spoke with Interactive Mechanics about the extent to which they had considered accessibility already in their design, and found that specific measures to improve accessibility had either already been built into the site or were on the to-do list. In particular, they had already assessed the site for high contrast throughout for readability and color contrast for the colorblind; proper use of structure and markup, titles, labels, alt tags, and title tags; and proper page structure and markup for screen readers. We recognized that there was a limit to the amount of budget money that we could allocate to accessibility during this first phase while our priority was still getting the site up and running as a prototype, and we agreed to reassess and devote more resources to Section 508 compliance after the site launch.

Testing design: after conversations with the institutional review boards at all four of our testing group's institutions, it was decided that IRB approval was not necessary as a prerequisite to the testing process. The testing group identified six main objectives, derived from core goals of the site, that would be served by usability testing:

- 1. The site must make its mission and main functionality clear from the homepage, without having to click down too many times.
- 2. An interested user who wants to know more about technical information, how the site was developed and funded, or where to go for further information should be able to find that very easily (one or two clicks).
- 3. The site should make it very easy for a user to quickly search for materials by keyword: it should serve users who want to explore primary sources about women's educational history without a particular project or research goal in mind.
- 4. The site should also serve researchers who are pursuing a specific goal, and allow them to guickly figure out how to do a more nuanced or advanced search.
- 5. The site should suggest meaningful connections between items at different institutions that are linked by common subject terms or themes.
- The site should make it possible for the user to pursue further research using the digitized materials, either by allowing downloads of the material, providing useful citations, or making possible communication with the partner institutions to get more information.

We wrote a <u>script</u> that would evaluate users' experience of the site with these objectives in mind. Each user was asked to sign a release form giving permission to record audio from the session, which was later used to flesh out the notes that contributed to our recommendations. The script prompted users to explore the homepage at their own pace, reflecting aloud as they did so, before attempting to complete a series of "tasks" designed to test whether certain key functions of the site were intuitively designed. In addition to a success/failure score, each session collected text from the users' comments on the aspects of the user experience that were positive or frustrating for them.

Testing: we conducted a series of scripted sessions with individual users, a testing facilitator from the team, and sometimes an additional notetaker to help capture feedback for the official results document. We reached out to staff, faculty, and students at and beyond the four institutions where team members were based, attempting to find a balance between targeting users whose perspective would be similar to the imagined audience for the portal project, and surveying widely to capture the broadest range of feedback possible. The eighteen testing participants included five staff (four of whom were librarians or education-technology affiliated), two faculty members, six graduate students (two of whom were in programs to become high school teachers), and five undergraduates. Six users were tested at Bryn Mawr, six at Barnard, four at Mount Holyoke, and two at Radcliffe. At least one participant was not in any way affiliated with the partner institutions. Tests were conducted at all institutions concurrently between February 16th and February 20th, 2015.

Results analysis: test results were in the form of data on the percentage of tasks successfully completed and the recorded dialogues between test facilitator and participant, giving us both quantitative and qualitative measures of the site's

performance. Broadly, our analysis found that users were compelled by the site concept and design, but struggled with content discovery. Our main findings included:

- Users liked the visual styling of the site, which they found aesthetically pleasing and "contemporary."
- The purpose of the site was unclear from a glance at the homepage: users understood the topic, but didn't understand that they could search a collection.
- Item discovery (defined here as searching for and navigating results through the "Search Collection" keyword search, Advanced Search, and Browse features) needed the most work. Users found the plethora of search types with different entry points throughout the site to be confusing. Instead, they wanted one streamlined entry method and more clarity about what type of search they were running (for example, a keyword search versus a title search versus a browsable list of items). We also discovered a need for smoother functionality with the results display and running subsequent searches (refining the results list or navigating to a fresh search).
- Users consistently expressed a desire to see more contextual information. Item
 descriptions that mentioned certain traditions, locations, or proper nouns left
 users wanting to see links to further historical information to fill in knowledge
 gaps. Generally, they called for a higher degree of integration with external sites
 (such as Wikipedia for basic references) and curated content within the site (such
 as blog posts or digital exhibits) to guide the user through the history.
- We discovered a number of small problems that could be fixed with small
 adjustments to usability and. For example, users intuitively clicked on the main
 image on the record pages expecting it to either enlarge the image or lead them
 to another page, and they were confused when it did not lead anywhere.

Implementation of recommendations: In March 2015, the testing group submitted an analysis of the results to the rest of the team and to Interactive Mechanics, with a set of recommendations that covered a broad range of suggestions but focused mainly on the crucial area of improving search and browse functionality. Interactive Mechanics worked closely with testing group leader Evan McGonagill (Bryn Mawr) to implement the group's recommendations. Over many iterations, Interactive Mechanics proposed and built alternative designs and expansions to the existing functionality of the site that addressed most of the key areas of concern. Most importantly, the final pilot site includes streamlined search and browse functionality that makes it much clearer what the purpose of the site is and how to use it. Interactive Mechanics also fixed many of the small issues that were not of crucial significance individually, but have a cumulative impact of establishing the site as a far more professional product. Users who revisited the site after the new development reported (informally) that it was a vastly improved experience.

Some feedback that we received could not be addressed within the limited financial and temporal scope of this grant.

 Several of the desired fixes to search functionality, such as being able to filter for item format, and providing an And/All/Exact option for search text,

- were either outside the capabilities of Drupal or would require further negotiations from the metadata group to standardized the items.
- One large area to be set aside for later was the contextual information that users desired to see: many people had expressed interest in seeing more curated material that would frame the contents of the site within historical context, which will require significant conversation among the project team to determine what types and depth of contextual framing would be appropriate and whether it would be realistic to dedicate the amount of staff time necessary to research and write the material.
- Enhanced functionality for the more textually-focused items--letters, diaries, and scrapbooks--was also discussed, and Interactive Mechanics designed experimental models for displaying transcripts and PDFs of those items that will be reviewed and refined in future iterations of the site.
- After using the resources on WebAIM.com to evaluate accessibility, the
 team also determined that Interactive Mechanics had done adequate
 work meeting Section 508 compliance for this phase. However, more
 targeted testing could be performed in the future to further enhance
 accessibility for the visually impaired, particularly in the area of screen
 reader testing.

APPENDIX D PROJECT PARTICIPANTS

A. Institutions' Staff:

Barnard College:

- Norberg, Lisa, Dean of the Library and Academic Information Services
- O'Neill, Shannon. Associate Director of Archives and Special Collections
- Tenney, Martha. Digital Archivist

Bryn Mawr College

- Appel, Rachel. Digital Collections Librarian
- McGonagill, Evan, Assistant Director, Albert M. Greenfield Digital Center for the History of Women's Education, and Interim College Archivist
- Mercado, Monica. Director of The Albert M. Greenfield Digital Center for the History of Women's Education
- Pumroy, Eric. Associate Chief Information Officer and Seymour Adelman Director of Special Collections (project director)

Mount Holyoke College

- Fields, Leslie. Head of Archives and Special Collections
- Goldstein, Sarah. Director of Digital Assets and Preservation Services, Library, Information, and Technology Services
- Trujillo, Shaun. Digital Library Applications Manager

Schlesinger Library, Radcliffe Institute for Advanced Study, Harvard University

- Aloisio, Paula. Archivist & Metadata Specialist
- Benson, Amy. Librarian/Archivist for Digital Projects
- Weintraub, Jennifer. Digital Archivist/Librarian

Smith College

- Lanzi, Elisa. Director of Digital Strategies and Services
- Myers, Elizabeth. Director of Special Collections
- Young, Nanci. College Archivist

Vassar College

- DiPasquale. Joanna, Digital Projects Librarian
- Streett, Laura. Archivist

Wellesley College

- Callahan, Jane. Archivist, Wellesley College
- Graham, Ian. Director of Library Collections, Wellesley College
- Hart, Kara. Systems Librarian Library & Technology Services, Wellesley College

B. Advisory Board:

- Gruber Garvey, Ellen. Professor, English Department of New Jersey City University
- Horowitz, Helen Lefkowitz. Sydenham Clark Parsons Professor, emerita, Smith College
- Kelley, Mary. Ruth Bordin Collegiate Professor of History, American Culture, and Women's Studies, University of Michigan
- Mandell, Laura. Director of the Initiative for Digital Humanities, Media, and Culture and Professor of English, Texas A&M University
- Rowe, Katherine. Provost and Dean of the Faculty, and Sophia Smith Professor of English Language and Literature, Smith College
- Tucker, Susan. Archival consultant and former head of the Newcomb Archives and the Vorhoff Library, Tulane University