Marquette University
e-Publications@Marquette

Library Faculty Research and Publications   Library (Raynor Memorial Libraries)

5-1-2015

Becoming the Gothic Archive: From Digital Collection to Digital Humanities

Rose Fortier
Marquette University, rose.fortier@marquette.edu

Heather G. James
Marquette University, heather.james@marquette.edu

Becoming the Gothic Archive: From Digital Collection to Digital Humanities

Rose Fortier

Raynor Memorial Libraries, Marquette University
Milwaukee, WI

Heather G. James

Raynor Memorial Libraries, Marquette University
Milwaukee, WI

Abstract:

The Gothic Archive is the flagship digital humanities project for the Marquette University library. The project was birthed from a simple digital collection, and through the partnership of faculty and librarians, was transformed into something more. The core tenets of digital collection creation were adhered to in order to create a solid foundation upon which to build the Archive. The expertise of both groups and communication were key in the evolution of the collection, and in discovering and highlighting the relationships between the objects. This case study reviews the steps Marquette took in creating the collection and taking it to the level of digital humanities project.

Keywords: Chapbooks, Digital surrogate, Digitization, Hoeveler, Institutional repository, Marquette University, Metadata, OCR, Supplemental materials

Introduction

Certainly, support for Digital Humanities (DH) research has grown in the past decade with the launch of the National Endowment for the Humanities (NEH) Digital Humanities Initiative in 2006 and the transformation of the initiative to the Office of Digital Humanities in 2008. Panapacker (2009) called DH "the first 'next big thing' in a long time" (para. 1). Multiple DH centers are popping up around the world, and the NEH and the Arts and Humanities Research Council (AHRC) continue to offer grant funding for DH-centric projects. So the question for librarians and archivists becomes, "what about us?" As Vandegrift (2012) offered in his response to Kirschenbaum:

> Libraries have struggled to define their role in digital humanities, as the discussions around DH often resort to theoretical discourse or technical know-how. Arguably, however, because the library already functions as a interdisciplinary agent in the university, it is the central place where DH work can, should be and is being done. DH projects involve archival collections, copyright/fair use questions, information organization, emerging technologies and progressive ideas about the role of text(s) in society, all potential areas of expertise within the field of librarianship. (para. 2)

Libraries are natural partners for DH projects; much of what DH needs, libraries have and librarians do. According to Vandegrift and Varner (2013) "Libraries and the humanities have always had a great deal in common." (p. 67). So libraries at the outset have had an obvious role in Digital Humanities research (Ramsay, 2010; Sula, 2013). Providing access and increasing wide accessibility have been historic tenets of librarianship, and the shift to digital access as a primary method has only increased librarians' awareness and efforts in this area (Kamada, 2010; Vandegrift, 2012). As the primary maintainers of institutional repositories and digital archives, librarians are concerned with preservation and sustainability of digital initiatives and have experience with the challenges of obsolescence and migration (Cantara, 2006; Cole, 2002; Thomas, 2013; Kretzschmar, Potter, Warwick, and Singer, 2010). Additionally, librarians have honed skills of collection development and curation that allow for broad recall without sacrificing coherence. This is crucial for current researchers as,

"information overload is now a hazard of the humanist's job." (Little, 2011, p. 353).

At Marquette University, the Gothic Archive is an online collection of primary resources and contextual materials that is currently a pilot project with hopes of one day becoming a flagship. This digital humanities project began as a simple digital collection and grew significantly in its first year of being, but it now faces the ubiquitous challenges that all digital humanities research projects must conquer or side step in order to persevere. It needs to conquer challenges of funding and technological limitations as well as needs for personnel and technical expertise. At the same time this project offers the ideal mix of opportunities for creating a robust and exciting digital humanities collection that can forward the research and innovation of scholars focused in various areas including, but not limited to, Gothic literature. This progression from digital collection to digital humanities project has included growing pains that are likely familiar to other library - digital humanities partnerships, but the current state of the Archive and its future potential lead all those involved to call it a success.

Background: Developing the Gothic Archive

The Digital Programs Unit of Raynor Memorial Libraries is just barely out of its fledgling state, having recently been re-named from its more temporary sounding predecessor, Digital Projects. A librarian coordinator, Rose Fortier, leads the unit and oversees a staff of 2.5 full-time employees as well as a rotating crop of student workers. The unit's main responsibilities center around e-Publications@Marquette (e-Pubs), the university's institutional repository, established in 2008. The repository serves as a showcase for Marquette's research output. In addition to faculty research and publications, e-Pubs also provides access to graduate student theses and doctoral dissertations. Digital Programs is currently involved in the retrospective digitization of thousands of theses, dissertations, and Master's essays. The unit also provides support for Special Collections and Archives digitization efforts.

In 2012 Dr. Diane Hoeveler, a professor in the English Department at Marquette, approached Fortier regarding the possibility of mounting her collection of digital images of rare gothic chapbooks in

e-Pubs. In support of her research, she had amassed them over the years by traveling to various repositories and photographing the chapbooks while she studied them. The original materials were short, cheaply printed, paper-bound books from the late 18th and 19th centuries. The gothic chapbooks had been popular reading material for the newly literate lower classes of Europe, but very few chapbooks remain. Those that do are scattered across multiple collections, very few of which have been digitized, none systematically. Hoeveler's research had focused on the British chapbooks and their anti-Catholic propaganda, and she was interested in creating a repository to bring together her digital copies of these fragile and largely inaccessible materials.

Hoeveler initially contacted the library because of an admitted lack of technical expertise on her part, but her vision for the project was ambitious. Beyond simply creating a repository to bring together fragile materials in poor condition, Hoeveler and Fortier discussed the possible goals of obtaining grant funding to continue Hoeveler's research and to gather and digitize the remaining gothic chapbooks in existence. Hoeveler wanted the collection to take on the overall shape of her research, which is multi-faceted and delves into the historical rise of literacy, developments in printing technologies, and the social contexts of the Enlightenment and Victorian England. Both Hoeveler and Fortier agreed that a simple collection of well-described gothic chapbooks might have been an adequate resource for a small group of English literature scholars. However, the Gothic has significant popular culture impact, meaning the project would be of interest to more than just literature scholars or even academics. Gothic literature, which is primarily fiction, emerged in the mid-eighteenth century and remained a popular culture phenomenon through the nineteenth century. Its core elements of horror, Romanticism, melodrama, gloom, suspense, mystery, morality, and rationality vs. the supernatural continue to influence many modern works as well, as does the genre's penchant for medieval and exotic settings. In an ideal world, the Gothic Archive could become a venue for scholarly discussion on the importance of the Gothic in literature as well as various other disciplines. For some still today, digitization of rare or fragile materials for access only is sufficient, but with researchers' increased use and familiarity with technology and born digital items, a flat collection is generally no

longer enough (Hughes, 2011; Terras 2012; Lynch, 2002). Lynch (2002) commented:

> We're getting pretty good at digitizing material at scale...the research questions are less about how to do it at all and more about how to optimize - how to do it more efficiently or effectively, how to be sure that you've chosen the most appropriate strategies and technologies. (para. 17)

Cantara (2004) has added,

> "missing are the tools and technologies humanities scholars need to visualize, analyze, interpret, collate, and edit texts and images in order to publish secondary monographs, scholarly editions, and teaching materials based on these primary resources" (p. 167).

Hoeveler and Fortier envisioned a digital resource within which the chapbooks would be surrounded by other materials, both historical and newly created, to give them context and connectivity. By providing the chapbooks and linking them to supplemental materials on their importance, Hoeveler anticipated the archive would provide a fertile ground for discussion. Ideally these supplemental materials would not only describe the relationships and interconnectivity between chapbooks, but also describe how the chapbooks related to thematic elements within literature, to elements within the Gothic, and to place them in a wider historical context. These relationships are what truly define a Digital Humanities project and elevate it above the level of simple digital collection. In flat digital collections, users must determine the connections between objects and their wider contexts. In a DH project, the connections and relationships are developed and illuminated through the project itself.

With lofty goals in mind, Fortier first had to determine which goals were accomplishable within the current budget and technical realities facing the library. Initial setup required little input from Hoeveler; beyond her ideas for the collection, she was content to let the library make decisions on how to proceed with the technical creation of the Archive. Given the size of Hoeveler's collection of digital images, Fortier determined the first phase of the project would be the establishment of a base collection – this turned out to be twenty-seven gothic chapbooks completed in just under twelve months. With the

availability of the e-Pubs software, Digital Commons, and the budgetary constraints of a project with no dedicated funding, Fortier confirmed that the collection should be housed within the institutional repository. A working group was made up of Fortier, the Libraries' Metadata Librarian, and a representative from Libraries' IT. In this way, the practical execution of creating the collection followed the standard habits of so many libraries who have done similar projects in the past. As Terras (2012) outlined in her overview of digitization practices, digitization of primary resources for the purpose of online accessibility is generally the first necessary step in any DH project with the most common form of digitization being basic digital images, which serve as digital surrogates prior to advanced computational analysis or manipulation. Digital images serve well as digital surrogates for documents, photographs, and two-dimensional art, which are the bulk of primary source material for most DH projects (p. 54).

These digitization efforts generally take varied sources in a wide array of states of preservation or decomposition and attempt to bring them into a standard format in order to allow searching and data mining of text (Terras, 2012, p. 54). For these purposes, text images can either be manually transcribed or converted via Optical Character Recognition (OCR) software in order to be able to search the text of the source (Terras, 2012, pp. 47-8). The end goal of this process is to allow scholars to "create, represent, organize, analyse, and communicate scholarly content" (Rieger, 2010, Technological frames, para. 5). This was true again for the Gothic Archive so that after creating digital surrogates of the originals from Hoeveler's images, optimizing them, creating embedded text, describing them, and loading them into the content management system - Digital Commons, in this case - were the next steps. For the first phase of the project Fortier determined that having the texts OCR searchable would be necessary and likely expected by users of the collection. As Nyhan (2012) has argued:

> Unsworth has discussed: 'One of the many things you can do with computers is something that I would call humanities computing, in which the computer is used as a tool for modelling humanities data and our understanding of it, and that activity is entirely distinct from using the computer when it models the typewriter, or the telephone, or the phonograph, or

> any of the many other things it can be. (Unsworth 2002)' From this we can make the analogy that the act of digitizing a text does not necessarily make it a digital humanities text, by which I mean a digital text that supports research and/or is the result of research.(Unsworth as cited in Nyhan, p. 118)

The first metadata for the Gothic Archive was what might be expected for any book: title, author, and publisher. Digital Commons uses the Dublin Core metadata schema, and the flexibility the schema affords would turn out to be a boon for the project in later phases. Basic keywords were added to the item records, similar to a catalog's subject heading, and the Libraries' Metadata Librarian, Lynn Whittenberger, offered guidance on fields that would be appropriate for the chapbooks. Fortier consulted with Digital Commons support techs and Whittenberger on constructing the metadata for the chapbooks with an eye toward other material types that might be added to the collection in the future.

At her suggestion, the keyword field was re-purposed to allow for thematic grouping. This keyword field would eventually lead to the creation of the thesaurus. She also championed the need for technical metadata. Even though needing to track the original reproduction mechanism was not implemented for the pilot phase of the project, it was something to be aware of and to put in place for future phases of the project. As more materials are added to the Archive, and from different sources, tracking the institution of origin and the quality and mechanism of the reproduction will become very important. Technical metadata increases user and researcher confidence in the authority of the digital copy. It also makes it possible to recreate the digital analog of the original in cases where the digital analog may become corrupted or inaccessible. Whittenberger also suggested creating a field to link from the chapbook to other materials, as those materials were added to the project.

During a 2013 campus mini-conference on "Conversations across the Humanities" Hoeveler talked with Heather James, the English Literature liaison for Raynor Memorial Libraries, about her desire to see the collection continue to grow and her interest in applying for grant funding to support the project. Shortly thereafter, James came on the project as a consultant to the future development and guidance of the project and a conduit between the Humanities

side of the project and the library's digital side. Hoeveler began the creation of supplemental materials illustrating the connections between the chapbooks and describing their significance on a wider scope and brought interested graduate students into the project to help speed the creation of these materials for the base collection of chapbooks. Supplemental materials, generated by Hoeveler and her research assistants included a chapbook synopsis, a short discussion of the chapbook in historical context, a list of keywords, and a bibliography. These materials were loaded into the Gothic Archive with their own metadata and linked using the keyword structure developed over the course of the second year. By the end of the 2013-2014 academic year, the Gothic Archive held 27 chapbooks and supplemental material for each of them, fully text searchable and organized by a keyword structure that is itself a resource, the Glossary of the Gothic. Given the accomplishments of the first phase of the project and its growing usage stats - over 5,000 downloads of materials, Hoeveler, Fortier, and James felt confident in applying for grant funding from the National Endowment for the Humanities in order to expand the collection and further develop the Archive as a DH project and resource to users of many disciplines and levels of expertise.

Figure 1.

The Gothic Archive landing page, showing the different pieces of the project.

Major Challenges & Solutions

In creating the Gothic Archive as a digital collection and then beginning to expand it to a more fully realized DH project, there were multiple challenges that had to be overcome, many of which seem likely to occur in other small scale start up DH projects. The lack of dedicated funding and the need to make a vision reality within the constraints of given tools are the likely context of many DH projects outside of flagship enterprises like the Yale Modernism Lab or the Maryland Institute for Technology in the Humanities at University of Maryland.

The first problem to become apparent was that the quality of the digital surrogates provided by Hoeveler varied greatly. She originally photographed the chapbooks for her own use, and the idea of the Gothic Archive came after she had already captured the images.

Even though the materials were transformed into digital form, they were not done to any standards. Furthermore, many of the libraries and archives she visited would not allow her to manipulate their unique and fragile materials so she could capture a cleaner image of the chapbooks, also diminishing the quality of the surrogates. To combat issues of page curvature, motion blur, and shadows, the images were extensively re-worked using Photoshop. This step was crucial to allow the creation of text since OCR returns the best results when the digital image has high contrast, is clean, and the lines of text are as straight as possible. With less-than-optimal digital files and compounded by the issues inherent in running OCR on old type-faces and characters the OCR results have not always been adequate.

Figure 2.
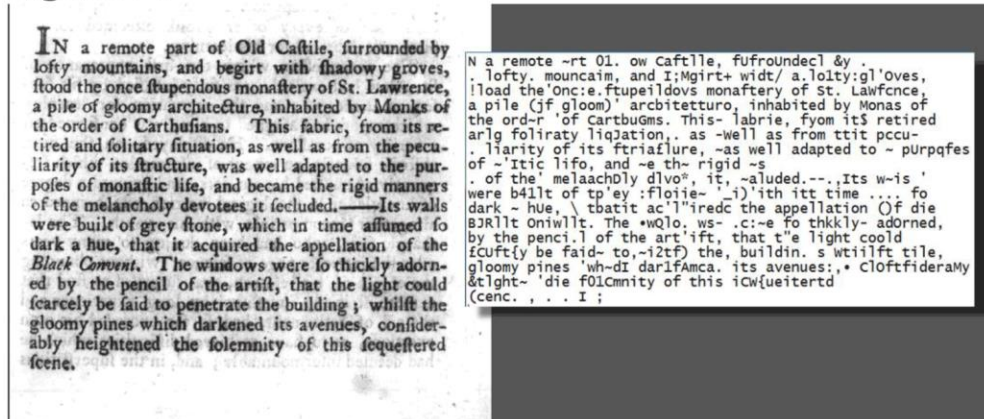
Side-by-side comparison of the original page and unedited OCR.

Re-working the images with Photoshop was time intensive, and contributed to the slow addition of chapbooks, especially at the beginning. Hoeveler had many more digital versions of chapbooks in her possession than ended up in the Archive, but even the best-quality versions required some re-working. Some were so poor in quality that the decision was made to transcribe the chapbook so that keyword searching via the full text would still be a viable option for the Archive's users. Perhaps the most important takeaway from this issue was that the chapbooks needed to be digitized according to current digitization standards, and that the institutions who owned the chapbooks needed to be involved. This is reflective of the development of many DH projects as being after the fact of scholars' original

research endeavors. It's tempting to see DH as an opportunity to simply throw Humanities research onto the Internet and share it with a wider audience, but projects that are not intended for DH from the beginning can offer plenty of problems later on. In future stages of the project Fortier and James plan to partner with the libraries owning the print chapbooks in order to replace the worst quality images with scans that meet NISO standards for digitizing.

After the problems with surrogate quality had been brought under control and the Archive started to be populated by digital chapbooks, the next hurdle was how to transform a flat collection into an interactive project while using software meant to act as a repository. The trend in user demands indicate an expectation for digital collections that also offer, "layers of interpretation and presentation built upon these databases and making reference to objects within them" (Lynch, 2002, para. 20). Though Digital Commons repository software was chosen for sound reasons, it was still first and foremost, repository software, best-suited for the static dissemination and preservation of materials. Additionally, the software is proprietary and tightly controlled by its developers. It became obvious very quickly that additional manipulation of the metadata was required to move the Gothic Archive beyond static repository. Somehow, the disparate pieces of the collection needed to be brought together into a multi-faceted whole.

Whittenberger was a particularly important player to consult in this area. Though the Gothic Archive is a self-contained project, it is housed within the larger institutional repository and has to remain consistent with already established local metadata standards. Metadata needed to be interoperable between the three main silos of digital collections and digital description at the libraries. These silos are e-Publications@Marquette, the digital collections of Special Collections & Archives, and MarqCat, the libraries' catalog. At her suggestion, keywords became the core of the structure to link the chapbooks with their supplemental materials, as well as a way to link thematic elements between chapbooks. In this area, the true partnership between librarians and faculty emerged. Once the first twenty-seven chapbooks were made available, Hoeveler oversaw the creation of supplemental materials by English graduate students. Included in the supplemental material were keywords that the graduate students
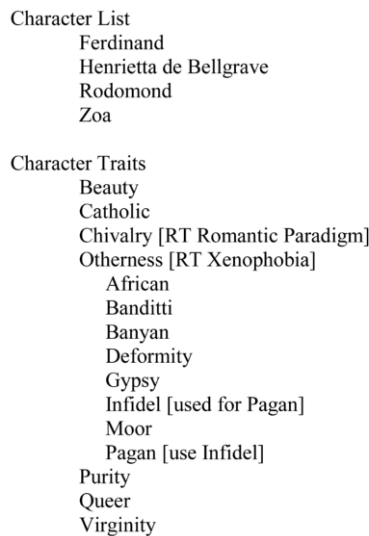
suggested for identifying key themes or content within the chapbooks. These keywords could be added to the internal metadata of the items within the Archive in order to link them. The question was how to apply that mechanism in the software and to ensure the tool was being consistently applied for maximum effectiveness.

The first part of the question was easiest to answer by seeking guidance from Digital Commons support. Among other things, support was able to offer a feature to auto-collect records into series based upon hooks in the metadata, allowing like chapbooks to be gathered together for browsing by theme. Creating hooks in the metadata underscored the importance of consistency in metadata creation and application, and once again Whittenberger was consulted to help determine a workable solution. She suggested using keywords as a way to draw together thematic elements in the Archive; a way of relating the chapbooks to one another. Fortier took it one step further, and proposed using those same keywords as a way of linking the supplemental materials to their chapbooks. The linkages had to work multiple ways. Thematic linking would allow the Gothic Archive to develop as a whole, and would allow the illustration of common themes and elements among the chapbooks. Functionally, the supplemental material needed links to their related chapbooks. However, they also needed to link to a central location where the controlled vocabulary terms could be collected and explained. This central location would serve as another way of developing the Archive into a cohesive unit, instead of a collection of individual works.

The next issue was the fact that terms were not being used consistently, and students were generating their own terms rather than working from a controlled vocabulary. This led to overlapping terms, synonyms, and terms unsuitable to the project. The project required specialized subject knowledge that could only be provided by Hoeveler and her students. At this point, James began working closely with Hoeveler's primary research assistant to pare down the keywords into a functional controlled vocabulary. Then Fortier and James continued working with Hoeveler and her research assistant to develop an index of terms with hierarchical structure, related terms, and preferred terms, including branches for themes, subjects, and Gothic Studies in order to effectively describe not only the primary sources but the supplemental material and future types of materials that may

be added. This index became the Glossary of the Gothic. The Glossary served two purposes: to deliver thematic background and context to the user and to provide a framework linking the individual pieces of the collection into a more visibly cohesive whole. Nearly every term has its own definition or description including scholarly references, so that the Glossary serves as an additional resource for the study of the Gothic in its own right. In this endeavor, the subject knowledge of the Humanities scholars married very well with the indexing expertise of the librarians.

Figure 3.

```
Character List
        Ferdinand
        Henrietta de Bellgrave
        Rodomond
        Zoa

Character Traits
        Beauty
        Catholic
        Chivalry [RT Romantic Paradigm]
        Otherness [RT Xenophobia]
            African
            Banditti
            Banyan
            Deformity
            Gypsy
            Infidel [used for Pagan]
            Moor
            Pagan [use Infidel]
        Purity
        Queer
        Virginity
```

Screenshot of part of the Glossary of the Gothic, showing the structure of the thesaurus.

Lack of familiarity with DH projects was also a challenge for all the librarians involved in the Gothic Archive. "Practitioners now recognize the potential of digital collections to function as components and building blocks that can be reused by many different groups and upon which many kinds of advanced digital library services may be built" (Cole, 2002, Starting premise, para. 3). How these collections might function and what various users might want from them can be large and even daunting questions. This was certainly the case for librarians and their introduction to Digital Humanities in this project. Fortier had been working with digital collections and digitization for over five years when she was approached by Hoeveler. However, she had neither the experience, nor the technical capacity to immediately move the collection to the higher levels of inter-relationships that are

the hallmarks of a DH project. She engaged in heavy consultation with multiple players in the library to tease out ways in which digital humanities differed from flat collections. James had very little experience with the management of the institutional repository and had to grasp the infrastructure of the software and its inherent limitations, and Whittenberger had no background with Gothic studies and its particularities. As the three learned and taught each other they were better able to communicate their capabilities and needs for the project and to develop a coherent plan with Hoeveler.

In the first few months, the timeline of the project was extended while librarians researched possible solutions to the challenges inherent in a project with such specific needs. Merely finding out if Hoeveler's goals were accomplishable took time while Fortier spoke with Digital Commons support, Libraries' IT, and Marquette University's Campus IT. Many of the most-used solutions for DH work today are open source. Unfortunately, open source solutions were out of the question because of Campus IT's refusal to allow the use of open source software. Even though many open source solutions are free or inexpensive, it was felt, by Campus IT, that such solutions would be a drain upon their resources. This inability to use open source software created a conundrum that has been difficult to overcome, and is still being struggled with as additional features to the Gothic Archive are being considered. The open source issue is even more frustrating as it would help with some of the financial limitations the librarians experienced.

The lack of DH experience also meant that there was a fair amount of need to discuss the definition and goals of Digital Humanities efforts in general and how the Gothic Archive might stake a claim among them. Research into what other institutions were doing informed the librarians' plans; in particular, they checked on the English Broadside Ballads Archive (EBBA) project at the University of California-Santa Barbara. Though there were software limitations and the Gothic Archive is at a much different phase in its development, the EBBA was an excellent starting point for an example of a similar effort. Ultimately, though it can be seen as disheartening, overcoming the lack of DH experience has left the librarians considering, as Thomas (2013) suggested, that due to the rapid advances in technology, we

can never effectively plan for an end product that is ultimately as valuable as we hope it might be:

> The distance between our wish and our object is often so great because the forms and practices and procedures of creation in the digital medium remain profoundly unstable and speculative...if a digital project becomes what was specified it might not be a digital humanities work. (p. 63)

Communication was less of an issue, and more of a constant effort to keep the project moving. There were multiple groups involved in the project at different points and in different ways. Librarians found themselves at the center of the need to communicate with subject experts (Hoeveler and her research assistants) and technology consultants (the Digital Commons support techs). Most central to the success of the project was the communication between the library and Hoeveler. Forging strong partnerships was essential in bringing the parties together; a common thread that runs through DH work. After all, Siemens, Cuningham, Duff, and Warwick (2011) have noted:

> Collaboration can enhance research by increasing its quality, depth, and scope and often can achieve what a single individual could not. However, these gains require additional work on the part of the team members and new skills such as conflict resolution, negotiation, communication. (p. 336)

Next to the grant writing process these collaboration efforts were most visible in discussion over the Glossary.

The academics provided the terms and explained their significance, while librarians organized them into functional hierarchies and reduced term duplication and overlap. The process required negotiation over terms and clarification over how the terms were used in the context of the chapbooks. Thesaurus creation underscored the differences in faculty and librarians' views. Faculty view the world through twin lenses of research and discipline. Librarians are more user-centric and take into account the needs of multiple user groups. Both points of view are valuable in DH projects, and satisfying the needs of both groups creates a more complete project. "It would appear that digital humanities projects and programs work best when interdisciplinary teams of librarians, IT specialists, and scholars come

together to share expertise, knowledge, skills, and perhaps most importantly, resources" (Little, 2011, p. 353).

This user-focused perspective is one that librarians bring, often uniquely, to a DH project as it develops. Librarians are aware from extensive experience that for whom a resource is meant or designed does not determine by whom it will be used. To their advantage, "a good academic library collection manager considers not only proximate core end users but also the wider and more diverse audience of both current and potential future end users, both near and far" (Cole, 2002, Starting premise, para. 3). For the Gothic Archive, as for many DH projects, the primary group may be the humanities researchers most closely interested in the primary sources, but as Lynch (2002) pointed out:

> I'm starting to believe that collections...don't really have natural communities around them...when you put materials out there, people you would never have expected find these materials from sometimes very strange and exotic places that you wouldn't have imagined, and sometimes make extraordinarily creative or unpredicted uses of that material (para. 34).

Finally, like so many start up DH projects, financial limitations were a challenge to overcome. Over the past couple of years, the Libraries have had budget reductions. As a result, new initiatives have had to happen without any additional funding. The Gothic Archive was not exempt from this unofficial mandate. The time and resources the project has taken have been carved out of already existing pools. For Fortier, this meant looking to use tools already at her disposal, one of the many reasons Digital Commons was chosen as the platform for the project. Fortunately, the Digital Programs Unit was already well-supplied with the basic hardware and software necessary for digitization, and Fortier and her staff were able to make do with those.

However, the major resource needed is time. The unit is a small one, and has many responsibilities. Juggling the already significant workload with the new project was not easy. Reliance upon well-trained students, especially at the graduate level was imperative. At any given time during the project, Hoeveler has had at least one graduate assistant working on the collection. Those students come with a level of subject and research knowledge that are vital. Because

of their schooling, they have not required the intensive time it would take to train another student or staff member. In the Digital Programs Unit, Fortier has made use of undergraduate student workers on various parts of the project. They have participated in aspects of the project that require a low level of training. Undergrads have been responsible for running OCR processes for the creation of embedded text. They have also been put to work on setting the automatic collection features in place.

Because the current phase of the project has essentially been a pilot project, Fortier worked her way through many parts of the more technical aspects of the process. Going forward, much more of the work will be delegated to students. Areas such as image processing and editing, and the uploading of chapbooks to the Archive are tasks for which undergrads are well-suited. However, without external funding, the project will advance at a slow pace. The current levels of funding for student workers allow Digital Programs to carry out the regular raft of responsibilities, but does not leave many extra hours for special projects, which the Archive is considered to be.

The issue of funding is a common concern in DH projects, particularly small scale start-ups. Terras (2012) has highlighted:

> Hughes details issues which may bring a sharp halt to digitization projects, including unresolved copyright issues, lack of adequate funding, lack of institutional support, technical drawbacks and the potential for digitization to damage or compromise fragile or rare original materials (2004, 50-2). (Hughes, 2004, as cited in Terras, p. 50)

Terras (2012) also said, "Whereas the 1990s were the 'decade of digitization' we are now in the decade of digital belt-tightening, self-reflection and honest assessment of achievements in using digitized content within the humanities" (p. 59). Funding is not as readily available as it once was, and demands are higher for the funding that is allocated to digitization projects today. One of the standard expectations today is consideration of long-term preservation and migration of the digital resources (Terras, 2012; Hughes, 2011). "Every digitization project… every funder of digitization projects...is acutely sensitive to this issue of sustainability" (Lynch, 2002, para. 24). And given the trend that many digitization projects are launched

with soft grant-based money but expected to be maintained on an ongoing basis, institutional commitment for the sustainability of these projects is necessary (Terras, 2012; Cole, 2002).

Yet, for all these significant challenges, the Gothic Archive has been a successful effort with promising potential. The collection has received over 11,000 hits and over 5,000 downloads of chapbooks and supplemental materials since December 2012. Though not yet widely known on campus, it has the potential to premier as the flagship DH project of Marquette University. This collection will hopefully bring more humanities faculty to the library to discuss their possible DH projects. It might also be the impetus to foster connections between humanities researchers currently siloed on campus. It has made the library realize that DH projects are large and complex. No one person (at least not in our budget realities) can be expected to tackle all that DH may contain, and so, at Marquette at least, this should truly be a group effort.

Beyond our campus the collection is significant to researchers and librarians for a variety of reasons. The chapbooks, freely available in their entirety, are also described in Marquette's library catalog and in WorldCat by extension, expanding the discoverability of these formerly inaccessible items around the globe. For libraries, each tale of small scale start up efforts can help confirm our efforts and support pitches for allocation of resources and personnel to make similar projects a reality. A recent American Council of Learned Societies (ACLS) report, as mentioned by Lynch (2012), describes libraries as the "laboratories of the humanities" (p. 467) which allow humanities researchers to do those elements of analysis and creation that Thomas, Unsworth, Cole, Lynch, Kirschenbaum and others celebrate. So a fully developed, or developing, DH project is significant in its contribution to the growing scholarly conversation. Even the smallest DH project must be more than a flat digital collection in its efforts to offer conversation between the materials and to allow future researchers to browse the materials and make serendipitous discoveries. In the end, the hope is that "the aggregation of materials in a digital library can be greater than the sum of its parts." (Lynch, 2002, para. 38)

Future Directions: Gathering the Remaining Chapbooks

The Gothic Archive, as it currently stands, is proof of concept. The work accomplished over the past year serves as a pilot project and has enabled the creation of best practices and workflows. With the completion of the two year-long pilot phase, the Gothic Archive was transformed from a digital collection of chapbooks and transcripts, into an interconnected DH project. However, the Archive still has a long way to go. As with any ambitious project, there is only so much to be done without funding. As Earhart (2011) has pointed out, "digital humanities is not a free venture, and in order to produce successful projects, institutions must provide some form of funding or support" (p. 33). Currently, Hoeveler is working with Fortier and James on grant applications for the addition of more chapbooks and supplemental materials to the Archive. Plans are in place to diversify the types of supplemental materials and include scholarly works on the Gothic and on chapbooks. However, the days of receiving funding for the simple creation of digital collections are past. Granting agencies are interested in partnership and innovation. Projects not involving significant aspects of both are unlikely to receive funding. To advance the Archive, funding is being sought to partner with other chapbook-owning institutions. By having those institutions digitize the items in their collection, the Archive will be expanded through high-quality digital reproductions of chapbooks, thus satisfying the need to partner, and the need for accurate digital surrogates.

Fortunately, because of the age of the materials, copyright is not a consideration for further developing the project. With the chapbooks ranging in publication date from the mid-1700s to the early 1800s, they are clearly in the public domain. It is important, however, to credit any institutions who might partner with the Archive through the contribution of digital copies of the chapbooks from their holdings. To address this issue, any such holdings within the Gothic Archive would attribute the owning institution. Attribution not only fixes the materials in further context, but also allows users to follow up with the owning institutions, if need be.

More materials exist that will further define the chapbooks' context and paint a fuller picture of their place in literary history and

the history of literacy. Hoeveler is gathering illustrations and frontispieces originally published with the chapbooks, for inclusion in the Archive. Full monographic works that illustrate overarching themes among the chapbooks are being digitized, and their importance will be discussed. In the long term, there are plans to create interactive maps to chart the spread of literacy among the lower classes and to use the maps to associate chapbooks to this spread.

Figure 4.

Example of a frontispiece for addition to the Gothic Archive.

More in-depth textual analysis and text-mining will be made available through the use of text-encoding initiative (TEI). This was not possible due to constraints of the Digital Commons software during the first phase of the project. Transcription and markup of the chapbooks using TEI will be time-consuming; the text will need to be proofread closely, and tags assigned. Some of that work has already been done; the transcripts mentioned earlier will be useful during the encoding process. Fortier will continue to work with Digital Commons support to encourage adoption of TEI capabilities.

As the Gothic Archive expands, new chapbooks will introduce new subjects and thematic elements to the Glossary, which will require the continued partnership between liaison librarian, faculty member, and research assistants. The expansion and population of the Glossary will lead to the creation of data visualizations, further illustrating the chapbooks' thematic connections. These further initiatives will possibly require the creation and migration of the Gothic Archive to new software. However, in continuing to use Digital Commons in the medium-term and ensuring the Archive is well-structured through the consistent use of thoughtful and robust metadata, the Archive will be well-positioned for migration.

Migration and software development are unlikely without more investment than can be offered through the Library or through the English Department. This kind of investment must be attained through grant funding, and from multiple grants. The Archive can continue to grow at a slow pace without this funding, though the most sophisticated tools and products are unlikely to be developed without external funding. The major cost in all this is the human cost. Purchasing hardware and software for such projects is easy, but the main investment comes back to wetware. People with the right skillsets and expertise are as necessary as equipment. The digital in digital humanities just denotes this era's tools for humanities research. The human element is as important now as it ever has been.

Conclusion

With the combination of Humanities research and digital tools, the key is partnership. The lessons of the Gothic Archive are that the project could not exist without the full cooperation and coordination of both main parties: the library and the faculty. Each group brings its specialties to the table. Libraries and librarians bring an understanding of digitization best practices which were applied to create the original digital collection. Further, their expertise in description and metadata serve to make the seed collection findable and interconnectable. Their considerations for searchability and broad accessibility across diverse user groups positioned the project to be seen, and used, by the broadest cross-section of users. Finally, their knowledge and concern for preservation techniques, both physical and digital ensure the Gothic Archive will persist, making it available to researchers, scholars, and other users for a long time to come. The Humanities scholars have

their roles to play, each as important as those offered by the librarians. Digital Humanities researchers bring a depth of subject knowledge that allows in-depth analysis of materials through which to discover and illustrate the connections that form the matrix of the DH project. They bring access to further subject experts through their networks and increase opportunities for collaboration. Neither group would have been able to bring the Gothic Archive to its current state alone. This project has been a true partnership complete with bumps in the road, with successes born of continued communication and perseverance that make for its own great story.

References

Cantara, L. (2004). Digital libraries in the humanities: Building a cyberinfrastructure for the humanities. OCLC Systems and Services, 20(4), 167-169.

Cantara, L. (2006). Long-term preservation of digital humanities scholarship. OCLC Systems & Services, 22(1), 38-42. doi:10.1108/10650750610640793

Cole, T. W. (2002). Creating a framework of guidance for building good digital collections. First Monday, 7(5), n.p. doi:10.5210/fm.v7i5.955

Earhart, A. E. (2011). Challenging gaps: Redesigning collaboration in the digital humanities. In A. E. Earhart, & A. Jewell (Eds.), The American literature scholar in the digital age (pp. 27-43) Ann Arbor, MI: U of Michigan Press.

Hughes, L. M. (2004). Digitizing collections: Strategic issues for the information manager. London: Facet Publishing.

Hughes, L. M. (2011). Evaluating and measuring the value, use and impact of digital collections. London: Facet Publishing.

Kamada, H. (2010). Digital humanities: Roles for libraries? College & Research Libraries News, 71(9), 484. Retrieved from http://crln.acrl.org/content/71/9/484.full

Kretzschmar, W. A., Jr., Potter, W. G., Warwick, C., & Singer, K. (2010). Library collaboration with large digital humanities projects. Literary and Linguistic Computing, 25(4), 439-445. doi:10/1093/llc/fqq022

Little, G. (2011). Managing technology: We are all digital humanists now. Journal of Academic Librarianship, 37(4), 352-354. doi: 10.1016/j.acalib.2011.04.023

Lynch, C. (2002). Digital collections, digital libraries and the digitization of cultural heritage information. First Monday, 7(5), n.p. doi: 10.5210/fm.v7i5.949

Lynch, C., & Carleton, D. E. (2012). Lecture: Impact of digital scholarship on research libraries. Journal of Library Administration, 52(6), 456. doi:10.1080/01930826.2012.707947

Nyhan, J. (2012). Text encoding and scholarly digital editions. In C. Warwick, M. Terras & J. Nyhan (Eds.), Digital humanities in practice (pp. 117-137). London: Facet Publishing.

Panapacker, W. (2009, December 28). The MLA and the digital humanities [Web log post]. Retrieved from http://chronicle.com/blogPost/The-MLAthe-Digital/19468/

Ramsay, S. (2010). Care of the soul. Retrieved from http://stephenramsay.us/text/2010/10/08/care-of-the-soul/

Rieger, O. Y. (2010). Framing digital humanities: The role of new media in humanities scholarship. First Monday, 15(10) doi:10.5210/fm.v15i10.3198

Siemens, L., Cunningham, R., Duff, W., & Warwick, C. (2011). A tale of two cities: Implications of the similarities and differences in collaborative approaches within the digital libraries and digital humanities communities. Literary and Linguistic Computing, 26(3), 335-348. doi:/10.1093/llc/fqr028

Sula, C. A. (2013). Digital humanities and libraries: A conceptual model. Journal of Library Administration, 53(1), 10-26. doi:10.1080/01930826.2013.756680

Terras, M. (2012). Digitization and digital resources in the humanities. In C. Warwick, M. Terras & J. Nyhan (Eds.), Digital humanities in practice (pp. 47-61). London: Facet Publishing.

Thomas, W. G. (2013). What we think we will build and what we build in digital humanities. Journal of Digital Humanities, 1(1), 63-67. Retrieved from http://journalofdigitalhumanities.org/1-1/what-we-think-we-will-build-and-what-we-build-in-digital-humanities-by-will-thomas/

Unsworth, J. (2002). What is humanities computing and what is not? Retrieved from http://computerphilologie.uni-muenchen.de/jg02/unsworth.html

Vandegrift, M. (2012, June 27). What is digital humanities and what's it doing in the library? In the Library with the Lead Pipe. Retrieved from http://www.inthelibrarywiththeleadpipe.org/2012/dhandthelib/

Vandegrift, M., & Varner, S. (2013). Evolving in common: Creating mutually supportive relationships between libraries and the digital humanities. Journal of Library Administration, 53(1), 67. doi:10.1080/01930826.2013.756699

Additional Reading

Adams, J. L., & Gunn, K. B. (2012). Digital humanities: Where to start. College & Research Libraries News, 73(9), 536.

Association for Information Science and Technology. (2013). DCMI metadata terms. Retrieved 4/29, 2014, from http://www.dublincore.org/documents/dcmi-terms/

Baca, M. (Ed.). (2008). Introduction to metadata, online edition (3rd ed.). Los Angeles: Getty Publications.

Bair, S., & Carlson, S. (2008). Where keywords fail: Using metadata to facilitate digital humanities scholarship. Journal of Library Metadata, 8(3), 249-262. doi:10.1080/19386380802398503

Besser, H. (2002). The next stage: Moving from isolated digital collections to interoperable digital libraries. First Monday, 7(6), 57 paragraphs.

Broughton, V., & Slavic, A. (2007). Building a faceted classification for the humanities: Principles and procedures. Journal of Documentation, 63(5), 727-754. doi:10.1108/00220410710827772

Bulger, M. E., Meyer, E. T., De la Flor, G., Terras, M., Wyatt, S., Jirotka, M., Madsen, C. M. (2011). Reinventing research? information practices in the humanities. Rochester, Rochester: Social Science Research Network. doi:10.2139/ssrn.1859267

Cornell University Library. (2007). Digital preservation management. Retrieved 4/29, 2014, from http://www.dpworkshop.org/dpm-eng/eng_index.html

Deegan, M., & Tanner, S. (2002). Digital futures: Strategies for the information age. London: Library Association Publishing.

dh+lib. Retrieved from http://acrl.ala.org/dh/

Federal Agencies Digitization Guidelines Initiative – Still Image Working Group. (2009). Digitization activities. (Technical Guideline). FADGI. Retrieved from http://www.digitizationguidelines.gov/guidelines/digitize-planning.html

Green, D. L., Albrecht, K., Cundiff, M., Friedland, L., Hirtle, P., Hughes, L., McKinney, P. (2002). The NINCH guide to good practice in the digital representation and management of cultural heritage materials. National Initiative for a Networked Cultural Heritage.

Guidelines: File format comparison projects still image and audio-visual working groups. Retrieved 4/29, 2014, from http://www.digitizationguidelines.gov/guidelines/File_format_compare.html

Hirtle, P. B. (January 3, 2014). Copyright term and the public domain in the united states. Retrieved 4/29, 2014, from http://copyright.cornell.edu/resources/publicdomain.cfm

Hughes, L. M. (2011). Evaluating and measuring the value, use and impact of digital collections. London: Facet Publishing.

Kirschenbaum, M. G. (2010). What is digital humanities and what's it doing in english departments? ADE Bulletin, 150, 55-61.

Lee, C. A. (. (2011). A framework for contextual information in digital collections. Journal of Documentation, 67(1), 95-143.

Martins, L., Helmer, S., & Arnold, D. Y. (2013). Exploring weaving structures in the andes: Reflections on the creation of a digital archive. Visual Resources: An International Journal of Documentation, 29(1), 59-71. doi:10.1080/01973762.2013.761121

McCarty, W. (2005). Humanities computing. New York: Palgrave.

Miller, S. J., & Cardinal, D. (2009). Metadata guidelines. Wisconsin Heritage Online.

National Information Standards Organization. (2007). A framework of guidance for building good digital collections. (guideline). Baltimore, MD: National Information Standards Organization. http://www.niso.org/publications/rp/framework3.pdf

OCLC Online Computer Library Center. (2010). Well-intentioned practice for putting digitized collections of unpublished materials online OCLC. http://www.oclc.org/content/dam/research/activities/rights/practice.pdf

Rimmer, J., Warwick, C., Blandford, A., Gow, J., & Buchanan, G. (2008). An examination of the physical and the digital qualities of humanities research. Information Processing & Management, 44(3), 1374-1392.

Sharma, R. N. (2003). The impact of digital collections on the humanities: A librarian's view. Against the Grain, 15(4), 48-52.

Sula, C. A. (2013). Digital humanities and libraries: A conceptual model. Journal of Library Administration, 53(1), 10-26. doi:10.1080/01930826.2013.756680

Svensson, P. (2009). Humanities computing as digital humanities. Digital Humanities Quarterly, 3(3)

Text Encoding Initiative Consortium. (4/18/2014). TEI: Text encoding initiative. Retrieved 4/29, 2014, from http://www.tei-c.org/index.xml

Unsworth, J. (2003). Tool-time, or 'haven't we been here already?'. Transforming Disciplines: Computer Science and the Humanities, Washington D.C.

Warwick, C., Terras, M., Galina, I., Huntington, P., & Pappa, N. (2008). Library and information resources and users of digital resources in the humanities. Program (London.1966), 42(1), 5-27. Retrieved from http://search.proquest.com/docview/62021956?accountid=100

Warwick, C., Terras, M. M., & Nyhan, J. (2012). Digital humanities in practice, London: Facet Publishing in association with UCL Centre for Digital Humanities, 2012.

Westcott, S., Bestebreurtje, L., Hoffman, S., Morton, A., Rhody, L. M., Schneider, B., Wieringa, J. Digital humanities now. Retrieved April 14 2014, 2014, from http://digitalhumanitiesnow.org/

Wisconsin Heritage Online Digital Imaging Working Group. (2009). Digital imaging guidelines. Wisconsin Heritage Online. Retrieved from http://recollectionwisconsin.org/wp-content/uploads/2013/01/WHO-Digital-Imaging-Guidelines-2009.pdf

Key Terms and Definitions

Chapbook – A form of early popular printing beginning circa the 16th century and continuing into the late 19th century. Chapbooks might include political, editorial, or literary content, but they are generally short, consisting of not more than 40 pages. They were generally printed on cheap paper en masse.

Digital collection – A group of digital objects in a content management system that are deliverable or accessible electronically. The objects are typically related in some way, whether by subject, theme, or format.

Digital surrogate – The electronic version of a physical object that is created through digital acquisition methods. The surrogate is usually as close to the original as possible and may have been optimized or altered so it more closely resembles the original.

Digitization – The process of creating a digital version of a physical object. The process starts with the acquisition of a digital image, usually by scanning, and progresses through other stages such as image optimization, and description.

Gothic – A literary movement beginning about the mid-1700s and continuing through the late 1800s. The Gothic genre is most well-known for its haunting novels, Frankenstein, Dracula, Jane Eyre, Wuthering Heights, many of which deal with motifs of death and solitude and themes of Romanticism vs. rationality.

Institutional repository – The online archive of an institution, usually a post-secondary institution. The archive typically includes faculty and student research and publications.

Supplemental materials – The texts, images, or other materials that add context to a digital collection but are not the primary sources of the collection. These materials may be contemporary to the primary sources of the collection or may be developed post hoc, and they may contextualize the primary sources within history or within a field of study or practice. They serve to further develop relationships between and significance of primary sources in a digital collection.