Building a Hosted Platform for Managing Monographic Source Materials and Born Digital Publications through Library/Press Collaboration

Summary

Michigan Publishing (home of the University of Michigan Press and a division of the University of Michigan Library), working in collaboration with colleagues at Indiana, Minnesota, Northwestern, and Penn State universities, intends to create a hosted platform for managing monographic source materials that is native to the modern collaborative infrastructure environment. The partners wish to request $898,912 in support from the Andrew W. Mellon Foundation to be spread over three years (April 1, 2015 to March 31, 2018). During this time existing Hydra/Fedora repository infrastructure will be extended to accommodate the ingest and interactive presentation of digital materials linked to humanities monographs through stable URLs and Digital Object Identifiers (DOIs) printed in paper versions and additional clickable links in electronic formats. Our project is enabled by the widespread implementation of XML workflows in university presses (that allow granular linking to materials outside the book and the creation of multiple versions of the work from the same source files) and the development by libraries of robust data repositories with capable access and preservation layers.

Our initial focus is on meeting the increasing demand university presses are seeing from their authors for ways of presenting source materials alongside interpretation. This need is often described by authors as being for “companion websites” although this is shorthand for a range of different requirements as explained by Jennifer Vinopal and Monica McCormick in their useful paper on "Supporting Digital Scholarship in Research Libraries: Scalability and Sustainability" (Journal of Library Administration, 53:27-42, 2013). Writing about the challenges facing New York University Libraries (to which the Press reports), they describe “the ‘faculty Web site problem,’ an ever-growing number of requests for Web-based spaces and tools to collaborate on scholarly research and share the results . . . , requests that actually represent a diverse set of activities which can be achieved in a variety of ways." The activity needs particularly addressed by the proposed project include the preservation of related data (e.g., further images, audio or video files), the deployment of open source tools that allow users to interact with source materials (e.g., searching a catalog of material, rendering of 3D images), and the delivery of data in a way that individual items can be cited and use measured so that authors get credit (e.g., through the use of Digital Object Identifiers (DOIs), through application of quantitative and qualitative usage metrics).

While there currently is some disciplinary resistance in the humanities to the presentation of entirely digital publications, we view the “companion website” concept as a culturally acceptable stepping stone to achieving such integrated presentations of narrative and data. In other words, while linking source materials to interpretation presented elsewhere is the main use case explored in this project, the proposed platform is also intended to be capable of supporting entire long-form enriched digital publications, complete in themselves. Rather than forcing them
to narrow their presentation to fit the constraints of a print (or electronic print facsimile) monograph, we hope to empower scholars whose research has been enriched through the application of digital tools to also publish their results in a rich electronic format, under the brand of a university press, and in an environment where preservation of datasets, the availability of tools to manipulate them, and systems to allow citation and measurement all coexist.

During the project, source materials for five humanities publications which illustrate a wide range of intellectual property, production, and preservation challenges will be made available through the system. Four of these publications present “data” as separate from “narrative” but in one case (Gabii) the authors have chosen to present an integrated digital publication as their only output. These five case studies will provide a proof of concept for an extensible hosted platform which, the partners intend, can then be offered to other presses on a fee-for-service basis. A feasibility analysis for this service will be conducted by a business planning expert.

We believe that our project is distinctive in the way that it not only builds on established technology but also leverages and advances strong existing relationships between university presses and libraries: All five presses are closely linked with their university library systems, four having direct reporting relationships, and this project will bring them even closer. Those libraries all share a commitment to the Hydra Project, are actively collaborating to build data repository infrastructure, and an Advisory Board of library/technology experts from the parent institutions will meet regularly throughout the project. At the helpful suggestion of the program officers, a series of collaboration and problem-solving meetings with other platform-development projects also invited by the Foundation to submit proposals to this grant initiative (led respectively by New York University Press and University of Minnesota Press) will also be held.

The Need
As humanists embrace the use of digital technologies in their research, the outputs of their work have become increasingly diverse and complex. An average archaeological project today, for example, produces thousands of digital images, hundreds of GIS maps and plans, multiple structured datasets, and many audiovisual files, in addition to the physical archives of notes and artifacts. Literary analyses increasingly employ text mining and analysis techniques that create a plethora of XML files; ethno-musicological studies may refer to thousands of specially-annotated music and film clips; theater texts may be illustrated with video clips of notable performances; and historical investigations may involve the creation of rich archives of digitized images of primary sources, richly tagged and annotated by the investigator. Despite the imaginative use of digital technologies in the process of research, however, and the rich records that are created, the publications that result from these projects predominantly retain the one-dimensional, text-heavy, non-interactive characteristics of traditional books, with digital versions essentially facsimiles of print.

In an informal survey of university press directors conducted on the AAUP-D list-serv in early November 2014, questions were asked about the level of demand university presses were seeing from authors for assistance in presenting digital source materials, the types of solutions
being offered, and whether there was a demand for the sort of hosted service being proposed. There was a split in responses between presses experiencing increasing demand and those that were not. Follow-up phone calls to some of the directors that responded suggested that two major factors at the presses that were experiencing demand were that the sub-disciplines with strong lists were ahead in embracing digital technologies (e.g., music, folklore, performance, archaeology, some areas of medieval studies and classics) and that the acquisitions editors were interested in digital scholarship and were willing to engage authors in discussions that went “beyond the book.” These respondents also indicated that they struggled with meeting the digital scholarship support needs that authors expressed, and would welcome solutions of the sort outlined in this proposal.

Here are some illustrative responses to the survey:

- “Your question about authors coming to us with digital components to support the text does resonate. But because of limited resources, i.e., staff and money, it is challenging to meet these demands. We do our best to accommodate manageable requests, like hosting on our website ancillary materials. It’s not always a smooth transition, though. Other authors come to us asking if we could help create an interactive site with video and links to supporting documents. This becomes much more difficult and we have to tell them 'no' or that they will need funding to hire a programmer."

- “To your point, we face situations like this all the time. Typically, the author has a methodological appendix that we won’t print, or some sort of supplemental material. We usually ask her to keep it on her own website rather than hosting it on ours. This is not an ideal solution in terms of archiving. I would say we face this situation about three times a year, but the examples are increasing, not decreasing.”

- “Yes. Most requests are for sites that will host musical examples with a few that are also interested in video. In most cases the music and the videos were recorded by the author as part of his or her fieldwork and are directly linked to the book. I am encouraging them to host the material themselves and explain all the ways we might link to their website in the book, on our website, and elsewhere. We of course do not have the technical staff to do these websites ourselves. I believe I get at least 4 or 5 queries about this issue at the music meetings, some queries at the folklore meeting, and in film studies we get a request or two a year for video clips. We always tell the author to work with his/her U to set up a separate website. So I and our editor in chief would welcome this sort of platform (with the ed in chief wanting to make sure it never goes away).”

- “We have certainly had increased demand over the past few years, mainly from disciplines where there are obvious needs/benefits to multimedia (film/cinema/media, music, etc), but increasingly some from social sciences (sociology, anthropology) where ethnographic studies may have video etc. available. The idea of just posting this on our website with a URL is a very clunky solution. And then there are issues about preservation that we haven't figured out.”

- “We do have plenty of authors who develop digital companions, websites, and archives for their books, several of which have been created in Scalar. And, yes, those numbers are increasing. At first these projects originated largely from media studies, but now we
have authors in performance studies, Latin American studies, and music developing such sites. Thus far we’ve asked authors to locate a permanent home for these digital materials before we will make any substantial links between our published book and the digital project. We’ll only do that if authors have a longevity plan in place for the digital site. For the most part that support has come through the authors’ own universities but the risk for authors is that they may move from institution to institution, and they would either have to leave the site with one university hoping that they will maintain it or move it to another university and risk disruption in access. What has been tricky for us are requests to host small bits of materials, such as extended bibliographies, extended essays, or ancillary materials. We don’t have a good way to store these items long term, and they are generally not substantial enough to warrant library preservation. In those cases, we are still seeking out good solutions.”

● “We have seen an increase in requests from every more authors regarding their wish to post source materials related to their publications. These cut across a wide range of subject areas and different types of source materials. Book authors often want to us to post data tables or supplementary material to support what is in their books. We do post this material on our website, open access, through Books Marketing. This is especially helpful for materials that support textbooks. When book authors want us to post material in proprietary application software, we are challenged and do not always accommodate. In general we believe we have been able to accommodate most editor or author requests with the resources available to our Press. The requests for data hosting are on the increase, however. If this should become an area where particularly new and complex file types or formatting were required it might make sense for us to work with other entities to address the need.”

Efforts to present source materials alongside monographic publications produced by university presses so far have either been one-off experiments or have lacked some of the essential characteristics of a sustainable system. The current solutions fit into several categories:

● **Links to author websites:** Authors sometimes ask for links to personal websites they maintain, but these are rarely stable enough to be suitable. At University of Minnesota Press the author of *Thirty Rooms to Hide In* includes home movies, photos, letters, and music in the companion website to his autobiographical book. The companion website is only linked to from the Press website, however, not from within the book due to concerns about permanence. At Purdue University Press books such as *Composing the Party Line* by David G. Tompkins use Digital Object Identifiers (DOIs) inside the book to reference the author’s personal website in the hope that should the hosting server move, the URL that the DOI links to can be updated. This of course presumes that a new location would be found.

● **Links to existing press websites:** For simple source materials, press websites designed for providing publicity about the publisher and its products are a solution also for mounting companion materials. For example, the University of Michigan Press traditionally has presented a selection of supplementary materials for books through web pages that are preserved by the library. But these are narratively and technologically
disconnected from the text, and do not easily facilitate citation of specific, individual source materials at URLs that remain stable over time.

- **Links to institutional repositories:** Supplemental materials can be presented in the institutional repository (IR) of an author’s parent institution but they are generally delivered in a “flat” way that uncouples the source files from the tools needed to manipulate them. For example, rich GIS data in Chapman University’s IR supporting a monograph about the spread of Greek vases around the ancient world are presented in PDF form or in files only usable with proprietary software. While some presses, like University of Akron Press, are able to place supplementary content from non-affiliated scholars in the institution’s repository, this approach sometimes collides with press’s parent IR’s policy and is done on a “special favor” basis that would be challenged should files be larger or more complicated.

- **Links to digital humanities centers:** When an author is fortunate enough to be at an institution where there is a center dedicated to supporting digital scholarship, sustainable companion websites can be established, but such scholars are in a minority. Notable one-off projects linked to university press books include the Dangerous Citizens Online project which involved collaboration between Fordham University Press and Columbia University Library’s Center for Digital Research and Scholarship (CDRS), the Guantánamo Lawyers Digital Archive created by New York University Libraries that is linked to a New York University Press volume, and the Civil War Washington website created by the Center for Digital Research in the Humanities linked to the book of the same name by Susan Lawrence, published by University of Nebraska Press.

- **Links to collaborative platforms:** In some cases, presses have built their own digital scholarship platforms, but have struggled to support them without a continuing flow of grant funding. A Mellon-funded initiative to create an Ethnomusicology Multimedia platform involved collaboration between Indiana University Press and Temple University Press but has suffered from a lack of early engagement with the library around long term sustainability. At the University of Virginia Press, the Rotunda project now supports several important collections of Presidential materials as well as the Society of Architectural Historians’ Archipedia. While there is a substantial subscription base, the platform still relies on a stream of grant funding to maintain its services. While platforms that aggregate content from multiple presses, such as Project Muse or JSTOR, have discussed providing access to supplementary content for books, these services have not been provided yet and there is concern in these organizations about how such customized work would fit with systems and processes that rely on standardization to keep administrative costs low.

It is clear from the above environmental scan that most current solutions are unsatisfactory. In an environment where all five partner presses and a number of colleagues in the university press community are increasingly being approached by academic authors interested in connecting their data to interpretation and creating more multimodal publications, the need is clear for a more systematic and regularized set of workflows and technologies. The opportunity explored here is to leverage library and university press competencies to facilitate the creation, maintenance, and presentation of publications in the humanities that begin to fully reflect the
richness of the research process. Even if our focus initially preserves the monograph in its familiar form (with four of the five case studies producing print versions), we believe that the platform we are creating will facilitate the creation of entirely digital projects published by university presses in the future -- perhaps even replacements for the monograph as presently conceived.

Substantially thanks to the encouragement of the Andrew W. Mellon Foundation, some other initiatives are now underway or being conceptualized to address the challenge of developing better platforms for the publication of digital scholarship. There are experimental tools that empower authors to write directly for readers, such as Scalar, “the born-digital, open source, media-rich authoring and publishing platform that’s as easy as blogging” (scalar.usc.edu/scalar). There are also redesigned forms of publisher-created platforms already used by multiple organizations such as the projected MUSE Open (currently applying for support from the Foundation), an extension of the Project Muse platform designed to allow Open Access book content to be presented in HTML 5 format, or by single companies such as a recently announced University of California Press initiatives to support open access books in partnership with Ubiquity Press and implement new workflows in collaboration with the California Digital Library.

Two initiatives particularly closely aligned with the project presented in this proposal are currently being also presented to the Mellon Foundation: These are the collaboration between University of Minnesota Press and the GC Digital Scholarship Lab at the City University of New York “Developing the Iterative Scholarly Monograph” and New York University’s press and library project around “Creating the Architecture for Networked Monographs.” While both projects aim to create platforms as end products, the Minnesota-led initiative primarily emphasizes the changed editorial procedures needed at any early stage in the production process to publish long-form digital scholarship while NYU’s focus is on the potential of semantic enrichment for discoverability once this work is published.

The partners in this grant proposal intend to remain closely connected with all these other projects. With NYU and Minnesota, whose projects are especially well aligned, we have prepared a plan for regular collaboration meetings which is described below. Given the diversity of disciplines university presses service, the existence of multiple platforms in the new ecosystem of scholarly communication is clearly desirable, and the projects at Minnesota and NYU are especially complementary in that (1) they explore different aspects of the publication process, (2) they are focused on the narrative rather than source materials, (3) they employ different underlying technologies.
Schedule of Major Activities

Design of the Platform

The proposed project focuses on the creation of a multi-tenant, cloud native, cross-branded platform for managing monographic source materials built on top of a robust infrastructure environment. It will feature an abstraction layer for enterprise workflow that will allow for preservation of monographic and similar long-form digital narratives, long-term sustainability and access for associated data sets, and long-lived access for media content that exists outside of the core monographic content within a networked global environment (on disciplinary repositories, in DPLA, on Project Muse, and JSTOR for example).

These broad requirements will be addressed by technology and system architecture choices designed to leverage existing and emerging library infrastructures common to the partner institutions. The basic repository and preservation system will be implemented in Fedora, with a number of Hydra heads to fulfill the publishing workflow and content access functions. This will form part of, and leverage, Michigan’s development of Fedora as an enterprise repository solution with Hydra heads for a variety of content types and workflows. The scope of Michigan’s Hydra/Fedora investment is illustrated below. In this diagram, which should be read from the center outward, the expressed needs of the University of Michigan Library Community (shown in the center) shape the Hydra heads required (light blue ovals in a concentric ring), which provide the services shown in the colorful “molecules” around the outside of the diagram. Existing work done is shown in blue text, while development needs are shown in maroon. The main focus of this project is to serve the Publishing needs shown in yellow to the left of the diagram.
Existing, publicly-available Hydra heads will form the basis of Michigan’s implementation when possible, with new features contributed back to the project or local customizations implemented as appropriate. Of the Hydra heads depicted, the following would support the proposed project, with existing heads that may form the basis of each indicated:

- **The Publishing Workflow** Hydra head, development of which is conditional on this proposal being funded, will provide the main interface for end-user access to the pilot projects. It will also support the submission, metadata creation, and access policy definition of source material content and route content to other appropriate heads based on content type. An existing Hydra head designed for the presentation of exhibitions (Spotlight) may serve as a basis for development. The Publishing Workflow head’s critical design features include:
  - **Branding**: It must be able to support unique branding for each publisher template, including providing a distinct Open URL for each partner case study and using Search Engine Optimization (SEO) best practices to ensure correct representation in discovery services.
- **Authentication**: It must have the capability to restrict access to content, potentially behind a paywall, as well as providing open access.

- **Multi-part publication packaging**: It must include a framework to assemble media assets (text, images, video, etc.) into a single “volume”. The media will be served from various Hydra heads designed to deliver their particular format (see below), with viewers embedded in the Publishing Workflow interface.

- A **Management** (MGMT) head will provide tools for governing the organization and presentation of the publication and configuring the branded interface for each participating Press, while also providing an overview of repository objects associated with each Press. Stanford’s Argo repository management and reporting module, which focuses on managing the integrity and curation of the whole Fedora environment, may provide a starting point for development work.

- An **Institutional Repository** head will provide data publishing, sharing, and curation oriented to single, stable objects, as well as persistent linking to both locally hosted content and external sources that have persistent identifiers. This head will build on Sufia, developed at Penn State University as a Hydra implementation of an institutional repository submission and retrieval system.

- A **Research Data** head extends the functionality of an institutional repository by supporting multiple file types with numerous relationships to each other (e.g., a code book, a “raw” data file with data captured from scientific instruments, and the Excel sheet containing cleaned and processed “cooked” data) that need to be modeled and stored as one complex object; different workflows for data capture and re-use that are used by different disciplines, that need to be accommodated by multiple workflows and multiple forms within the research data service; data sharing use cases that need to accommodate "dynamic" data sets, e.g., data sets within the repository that are continuously being updated as new research is conducted on the data set, and require tools for managing such things as provenance and more complex curation decisions for long-term preservation; and integration points into campus computing infrastructure, such as high performance computing clusters. Research data also has a compelling need for more robust inter-institutional data sharing and discovery services to plug into regional- and national-scale repositories; see for example the National Data Service initiative (http://www.nationaldataservice.org/).

- **Image**, built on the Digital Image Library head (DIL) from Northwestern University for collections of still images.

- **Video/Audio**, built on the Avalon Hydra head from Northwestern and Indiana, for support of time-based media. Both the Image and Video/Audio heads will provide embeddable players for viewing and interaction from within the context of the assembled volume in the Publishing Workflow head. It is worth noting that we have received a commitment from the Avalon development team, recently announced as recipient of a further support from the Mellon Foundation, to full cooperation in integrating Avalon into this project.

- **A Reserves** head will implement a number of functions specific to managing and providing access to library reserves, yet will be repurposeable for scholarly publishing:
setting granular permissions, allowing a mix of publicly accessible items with items restricted to individual or institutional access;

○ content reuse and repurposing through different channels (e.g., via the Unizin digital learning platform);

○ a specific form of use tracking that meets requirements for learning analytics (e.g., providing information about who uploaded the item, what classes have reused the same content over time, what semesters the content has been used in, which students have viewed it, when the file was last changed and who did it, and various other measures that relate specifically to class-based, copyright-restricted viewing).

Other applications, not shown in the diagram but listed below, will likely be necessary for the full functionality of the hosted platform to be achieved. The first two are essential for the presentation of fully-integrated publications where narrative and data are presented together. The third application, a 3D Scene Viewer, is most relevant to the Gabii project and is being developed primarily with other grant support.

- **Blacklight** will provide the ability for users of the companion websites to conduct full-text search of any texts.

- An **XML Text Renderer** to perform the transformations from TEI or JATS XML to HTML for viewing in the browser. Existing XSLT libraries, including the XSLT used by Michigan’s own DLXS platform, can be used or adapted for this purpose.

- A **3D Scene Viewer** for rendering 3D georeferenced and realistic models in the browser. The Gabii Project is seeking other sources of funding to continue design work on a generalizable, fully interactive 3D scene viewer, potentially in collaboration with other Hydra partners. As well as funding in place from a current (2013-2015) NEH ODH Startup Grant (*The 21st c. Data, 21st c. Publications. 3D Model Publication and Building the Peer Reviewer Community Project*), the Gabii project plans to carry out further work on the 3D modeling through an NEH ODH Implementation grant (application deadline February 2015), building upon their successful startup grant. This work will leverage the planned WebGL output option for the Unity3D gaming engine (Unity 5 release; currently in beta, full release 2015) and include an improved user interface and added functionality. The current version of the Gabii Project's 3D Scene Viewer is operational, and runs in browser with the help of the free Unity3D plugin. Until the WebGL version is available in 2015, the plugin dependent version will be used, with a pre-rendered static snapshot available for users with restrictions on loading plugins.

**Identifiers:** To facilitate persistent access to content and interoperability of systems, the hosting service will also include the registration of Digital Object Identifiers (DOIs) for each discrete piece of content deposited into the system. Leveraging Michigan’s membership with CrossRef in this manner supports the durability of access to the content, increasing both its citability and, by extension, its discoverability. RESTful APIs will be deployed to allow machine interaction with the service. The use of CrossRef DOIs (as opposed to other stable identification schemes) also has the advantage of firmly situating the published materials within the traditional publishing
environment, with a number of authors now associating CrossRef identifiers with “trusted content.”

**Analytics:** We believe that rich usage statistics are essential for driving acceptance of, and enthusiasm about, the publication of digital scholarship. We will provide quantitative and qualitative analytics to authors, publishers, and sponsors in several ways:

- An analytics dashboard viewable within the platform that will display basic usage statistics collected via the Google Analytics API, including visits, pageviews, pages/visit, bounce rate, average time on site, % of new visits, top pages, top referrers, and top searches;
- A monthly report sent via email showing the basic usage statistics available in the dashboard; and
- Access to the underlying Google Analytics account (separate interactive views for each Press) which makes available the full range of metrics Google Analytics collects, including data on technology used (browsers, operating systems, mobile/desktop), traffic channels (search, referrals, direct, social media), behavior flows, and visitor locations (e.g., country, continent, city).
- We are also pursuing a relationship with Altmetric.com which would allow us to provide additional insights into the wider impact of these research publications to authors, publishers, and sponsors by tracking social media mentions, reports in news outlets, and other conversations about the work happening on the open web.

**User Privacy:** Data collection practices will conform to the User Privacy Policy of the University of Michigan Library. This clearly articulates how usage and user data would be collected in ways that protect user privacy. All partner presses have agreed that using this policy for the hosted platform would be acceptable to them, but explicit agreement to do so is still required between the Office of the General Counsel (OGC) at Michigan and its counterparts in each partner institution. As expressed in the signed Shared Statement of Intent documents in the appendices, formal agreements will need to be signed between the partner institutions if the proposal is successful, and acceptance of the privacy policy will be one part of these. Attorney Jack Bernard in the Michigan OGC, also the lead author of the Michigan Library User Privacy Policy, is poised to working with his CIC colleagues to reach such mutually acceptable agreements should our proposal be successful.

In the case of content held behind a paywall, payment processing and customer service will be the responsibility of the respective publisher (e.g., Indiana University Press) following the terms set in place by their parent institution for financial recordkeeping. At Michigan, where credit card processing is handled by Nelnet, all staff involved are required to follow the Payment Card Industry (PCI) Data Security Standard. The consumer data transmitted to Michigan Publishing to allow authentication of each title or publisher collection on the platform will also be subject to the User Privacy Policy of the University of Michigan Library, which states that data matching user identities to materials accessed will be collected and retained “only for the purpose that such record is necessary to furnish a specific service” (in this case, to provide ongoing access.
to the content) and that “data on individuals will not be shared with third parties unless required by law.” A further layer of protection is also provided under Standard Practice Guide Information Security Policy of the University (SPG 601.27).

**Hosting Infrastructure:** In the interest of maintaining a consortial arrangement of offering services on shared infrastructure, the platform will be deployed as a cloud service. Michigan is assessing a variety of cloud services for different purposes: Amazon Web Services for storage and compute services; Microsoft Azure for computational power and specialized services such as large-scale file format migration; and VMware vCloud for burst computing, disaster recovery, and point-of-need infrastructure. Deployment will begin on local infrastructure, with decision on cloud infrastructure to be made by fall 2015, and the platform migrated into the cloud for final production deployment in Spring 2017.

The preservation of materials in the Fedora repository will be greatly enhanced by Michigan’s serving as a node in the Digital Preservation Network (DPN), by which dark copies of the material will be replicated among other DPN partners to protect against catastrophic failure and ensure the longevity of the data.

**Creation of Case Studies**

Central to the proposed project is the production of the five case study publications within the timeframe of the project. In the case of the projects from Indiana, Minnesota, Northwestern, and Penn State the interpretative portion of the publication will be made available in print and on conventional ebook platforms with the project described in this proposal providing the home for richly-presented source materials. The partner presses have varying ebook distribution relationships but generally use Amazon Kindle and to a lesser extent Apple iBook for direct-to-consumer (D2C) sales and JSTOR Books, Project Muse/UPCC, Ebrary, EBSCO Books, and Overdrive for business-to-business (B2B) organizational purchases. Connections will be through Open URLs and DOIs in the print book supplemented by clickable links in electronic versions, to the extent supported by different devices. In the case of the Michigan project, there will be no separation between the interpretative and the source materials, with the entire long-form digital publication only appearing on the Hydra Fedora platform. Final versions of the projects will be released at regular intervals throughout the second and third years of the grant, with the University of Michigan’s project being last due to the additional complexity of rendering 3D models that it involves. One of the first activities that the team will engage in is the abstraction of the central needs of all five projects and a limited number of templates will be developed which will be repurposed.

The fact that the project at Michigan is the only integrated presentation of data and narrative is mostly the result of what projects under or near contract were available for the partner presses to use as case studies. The Gabii team has already shown their commitment to enriched digital publication as the version of record during the peer review and selection process, which included evaluation of underlying data alongside narrative using workflows developed with funding from their NEH Office of Digital Humanities startup grant. As described in the introduction to this proposal, there is still substantial cultural resistance among humanists to the
idea of an integrated digital publication without a separate print component. The “companion websites” being produced for four of the case studies are stepping stones to acceptance, but only the authors of the Michigan project were willing to take the risk of waiving the requirement of a separate narrative presentation formatted for print.

There is a variability in perspective among the partners as to the extent of open access offered to the companion website. Minnesota, Northwestern, and Penn State envisage the source materials being openly available, but Indiana and Michigan expect at least some of the source materials to be access restricted and potentially behind a paywall. One of the primary design criteria for the Publishing Workflow head is that it have the ability to restrict access via Shibboleth authentication. While the financial transaction and customer service would need to take place outside the proposed Hydra Fedora system and would be the responsibility of the relevant publisher, access to content on the companion websites could be restricted to particular IP addresses (for organizations) and email addresses (for individuals).

A particular issue that the partners have discussed is the relative phasing of the release of the interpretative books and their source material collections, in the cases where the two are separate. Ideally both would be released simultaneously (or, in the future, the companion site might even be released earlier to generate “buzz” for the book). However, the experimental nature of the project means that such synchronization may not be possible. In this context, the University of Michigan has committed to minting the Open URL of the companion site and providing a temporary presentation of materials using the existing Sufia Hydra head to coincide with the publication of the book (a beta site), with the understanding that the full functionality may only become available later (a final or production site). This may also affect the partners’ strategies for releasing ebook versions, with a likely separate “enhanced ebook” release envisaged after the companion site is complete that would feature more granular linking and greater interoperability and functionality.

A simplified workflow for the partner presses, where the book and companion site are separately developed, is shown below. As described below, the main complexity lies in the synchronization of the two parallel processes. This workflow is inspired by the Dryad organization (http://datadryad.org/) who have mapped how article and data publication can be managed side-by-side in scientific journal publishing. While technologically more challenging due to requirement for 3D rendering, the Gabii workflow is in some ways simpler since there is no separation between “book” and “companion website.” These proceed in the same workflow from the point of selection onward. As previously mentioned, for Gabii the underlying data were subjected to peer review in the same way as the interpretative layer in a process organized by the University of Michigan Press. This was accomplished with the support of a 2013/2015 NEH Office of Digital Humanities startup grant supporting the project “21st c. Data, 21st c. Publications. 3D Model Publication and building the Peer Reviewer Community” (PI: Fred Limp, University of Arkansas).
Figure 1: Workflow when data and narrative separate
Indiana University Press - *Sustaining Place through Music: Performance as Environmental Activism*, by Mark Pedelty, University of Minnesota. Under contract, book publication projected for fall 2016 in simultaneous print and ebook editions. This book explores the vibrant ec-musical history and contemporary community of Cascadia (the towering mountains on the coast of Washington and British Columbia) in order to understand how environmentalist music imagines a more sustainable conception of place. Alongside encroaching developers and
industrialists is the presence of a rich environmental movement that has historically built
community through musical activism. From the Wobblies’ *Little Red Songbook* (1909) to Woody
Guthrie’s Columbia River songs (1941) on through to the Raging Grannies’ formation in 1987,
Cascadia’s ecology has inspired legions of songwriters and musicians to advocate for
preservation through music. Pedelty examines the divergent strategies (musical, organizational,
and technological) used by each musical group to reach different audiences and to mobilize
action. Through the proposed project, Mark Pedelty will be able to share his rich trove of
interviews, performance clips, and audio samples he has collected over time. Envisaging the
electronic supplement, the author has been rewriting sections of the manuscript to accompany
links in the text to particular multimedia files and the complete text of the book with these
additions will be presented in the companion site. While some of the linked material has been
published before, the interviews are original and rights releases have been obtained from the
subjects. The preservation commitment underlying the proposed Hydra Fedora platform is
particularly important in the case of these unique files.

**University of Michigan Press:** *A Mid-Republican House from Gabii*, edited by Rachel Opitz
(University of Arkansas), Marcello Mogetta (Free University, Berlin), and Nicola Terrenato
(University of Michigan). Accepted by editorial board, contract in the works. The Gabii Project is
a major archaeological excavation focused on a primarily Archaic site near Rome in central
Italy, conducted under the aegis of the University of Michigan. Since the start of the excavations
in 2009, over 500 GB of digital data have been collected during five summer fieldwork
campaigns in the form of photographs, scanned plans, surveyed measurements, written
descriptions, sketches and 3D models capturing stratigraphic layers and features such as walls,
pavements and burials. Having collected and created this rich digital dataset, the project
directors face the challenge of developing effective ways of publishing it. The intention is to
allow users to access the records related to the excavation of one of the sectors of the site,
Area B, in three ways: (a) through reading a hyperlinked narrative, (b) by manipulating a 3D
model, and (c) by accessing downloadable datasets. The “volume” is planned as an all-digital
publication with no print component. There has been substantial discussion already of the
additional complexity this project brings, and its development is also receiving support from
NEH and the Kelsey Museum of Archaeology at the University of Michigan.

**University of Minnesota Press:** *A Cultural History of the Canoe*, by Mark Neuzil, University of
St. Thomas, and Norman Sims, University of Massachusetts Amherst. Under contract, book
publication projected for spring 2016 in simultaneous print and ebook editions. Written by two
historians who are also experienced canoeists and canoe builders, *A Cultural History of the
Canoe* is a scholarly history of the canoe in North America and beyond; from its origins in the
Caribbean and South Seas, through adoption by Europeans who wished to travel into the
interior of North America, to its widespread use in the 20th century for leisure, recreation, and
further exploration. The paths of cultural transmission and the remarkable stability of the form of
the canoe are two themes that run through the book and make it a valuable case study in both
social history and the history of technology. The work includes archival photographs, drawings,
paintings, maps, and other visual records of the canoe’s history. As well as allowing versions of
these and additional images to be browsed and zoomed in on, the companion site facilitates the
presentation of multimedia elements such as historical film clips, explanatory videos showing boat construction, and audio excerpts. The site also permits the preservation of original video interviews undertaken by the authors with master boat builders and of third party materials that have previously been presented elsewhere on the web but in unstable ways. The particular challenge for the project with this book is that it is intended as a crossover academic/trade publication. The authors have backgrounds in journalism as well as history and sales expectations are high. This sets a higher level of expectation around the design and ease-of-use of the companion site since users will be from outside as well as inside the specialist academic community.

**Northwestern University Press:** *The Director's Prism: E. T. A. Hoffmann and Russian Modernist Directors*, by Dassia Possner, Northwestern University. Under contract, book publication projected for spring 2016 in simultaneous print and ebook editions. Possner’s book examines the creative work of four of Russia’s most significant avant-garde directors (Theodore Komisarjevsky, Vsevelod Meyerhold, Alexander Tairov, and Sergei Eisenstein) in the context of their fascination with German Romantic writer E. T. A. Hoffmann (1776–1822). From 1910 to 1922, when Hoffmann inspired a frenzy of activity that has been referred to as a “cult” of Hoffmann, these avant-garde directors transposed literary techniques from Hoffmann's stories and novellas to theatre and film, using them to refract, fragment, and otherwise artistically distort their world. In so doing, they held up a subjective, creative mirror to Russia’s chaotic Revolutionary period. The companion site presents digital versions of the notes and designs of these directors, archival photographs, sound recordings, film clips, librettos, and musical scores, many of which items are currently difficult-to-access in Russian archives.

A particular note on this project: The intellectual property aspects of this site will be particularly challenging as the Press will be dealing with Russian copyright legislation and a variety of overseas stakeholders. However, both the Press and the author have plans in place to minimize risk of failure. The author plans to work with six repositories in Russia: The Bakhrushin State Central Theatre Museum, Moscow; Hermitage Museum, Saint Petersburg; Russian Museum, Moscow; Russian State Archive of Literature and Art, Moscow; St. Petersburg State Museum of Theatre and Music; and Vakhtangov Theatre Museum, Moscow. Of these, NUP has published books that included illustrations sourced from three of them: the Hermitage, the Russian Museum, and the Russian State Archive – which holds the majority of material that the author plans to use. The author has previously obtained permission to use material from the St. Petersburg State Museum of Theatre and Music and has colleagues who have successfully reached permission agreements with the Bakhrushin State Central Theatre Museum. This leaves only the Vakhtangov Museum as the one repository about which relatively little is known regarding their practices. It should also be noted that the primary figures in the author’s study (Meyerhold, Tairov, and Eisenstein) left no heirs which will simplify matters. Taking all of these factors into consideration, it is highly unlikely that reaching agreements to use material from these repositories would be unusually difficult.

**The Pennsylvania State University Press:** *Nothing but Love in God’s Water: Vol. 1: Black Sacred Music from the Civil War to the Civil Rights Movement, and Vol. 2, Black Gospel Music*
from the Civil Rights Movement to the Present, by Robert Darden, Baylor University. Under contract. Volume 1 was published in fall 2014, and volume 2 is projected for fall 2016. Concurrent with the release of the second volume, the press will make both volumes available in digital formats with enhanced media links to the companion site embedded. This publication chronicles the history and role of music in the African American experience and explores how songs and singers helped African Americans challenge and overcome slavery, subjugation, and suppression. From the spirituals of southern fields and the ringing chords of black gospel to the protest songs that changed the landscape of labor and the cadences sung before dogs and water cannons in Birmingham, sacred song has stood center stage in the African American drama. Interviews conducted with performers will be preserved on the companion site as well as examples of early songs from the Alan Lomax collection, the Black Gospel Music Restoration Project at Baylor University (jointly run by the author), and images, sounds, and video clips of the Civil Rights Movement from the Library of Congress. A particular opportunity with this site will be to showcase our system’s capacity to link to external sources.

Sustainability

The main route by which the technology and workflows developed during this project will be sustained is through adoption and use by the Hydra community. Initiated by Stanford University, University of Virginia, University of Hull, and Fedora Commons in 2008 as a collaborative, open source effort, Hydra has had many years of self-funded productivity, and can demonstrate a growing code, contributor and user base. Over twenty-five official signatories to a partnership agreement are listed on its site, including a number of large research universities, but many other institutions are using some of the web-based applications, or heads, developed. A group of large research libraries including Michigan, Indiana, and Illinois have recently explicitly adopted Hydra as a solution for their data repository services and are adopting a shared approach to development. Michigan will use the Publishing Workflow head as a replacement for its aging DLXS publishing platform and thus will provide a continued development and maintenance commitment. We are also confident that other institutions that already deploy Hydra in other ways will also adopt the technology developed.

In addition to relying on the Hydra community for sustainability, the partners to this project will engage a consultant to develop a business model for a hosted platform offering for presses without access to the appropriate library infrastructure. The business model will be formulated in the first three to six months of the project, a period of time when the consultant will: (1) Conduct a structured fact-finding process (“discovery”); (2) develop a preliminary model to share with the project team; (3) refine the model before the build-out takes place. During and after the build-out, the model will be tested, confirmed or amended, as needed.

In conducting the structured fact-finding process, the consultant, working with members of the project team, will identify constituent and competitive groups for the platform. Constituent groups will include the presses and libraries participating in the project, other university presses that are likely to be customers of a hosted platform, and library publishers that may become users of the platform. Using a structured template, the consultant will interview 12 to 15
representatives from the presses and libraries identified by the project team or the consultant. A number of organizations of varying sizes and profiles have already expressed interest in participating including Temple University Press, Cornell University Press, Fordham University Press, University of Illinois Press, University of Nebraska Press, University of Mississippi Press, and Amherst College Press. The conversations will be documented and the information gathered will be used to develop a first-pass model. The consultant will also research and potentially interview competitive, substitute, or “good-enough” solutions for hosting content and providing adequate access. Three to five additional interviews or research calls may be made as part of this effort.

Using the data gathered as part of fact-finding, the consultant will then develop a preliminary model. This will present five-year revenue projections integrated with cost and timing information as best known at the time this preliminary model is created. This model will be shared with members of the project team, in all likelihood several times, so that the team gains an understanding of its structure and implications.

The model will then be refined before the build-out takes place. Options, changes and suggestions made by the project team as it becomes more familiar with the model will initially be made as they come up. After a reasonable period of time, the model will be “locked” and turned over to the team, which will be able to use it as a reference during the build-out and early test cases. In effect, the model will serve as a kind of financial and operating budget, against which the project team can measure its progress and its ability to gain customer support for a hosted platform.

At various check-points along the way, determined jointly by the project team and the consultant, the model itself will be revisited and potentially revised based on actual experience. With the initial model set at six months, these check-in evaluations will take place at the end of years one, two and three.

The deliverables of the consultancy will be: Documentation of the fact-finding discussions, with data made anonymous where appropriate; a competitive assessment; a preliminary business model; a “build-out” version of the business model; refinements as needed during the course of the project.

Expected Outcomes and Benefits

The outcomes of the project will be: (1) a proof-of-concept system and business plan for a fee-for-service hosted solution for monographic source materials; (2) the publication of an open source “solution bundle” for institutions using Hydra/Fedora who prefer to develop a local hosting option; (3) the development of a workflow and relationship model for presses and libraries working together to create enriched humanities publications that leverage their complementary strengths and systems, exemplified by five real publications.
The initiative leverages the new kinds of relationship that have been developing between an increasing number of university presses and their parent institutions’ library systems. Over 25% of AAUP members who self-identify as “university presses” now report to library directors, and those relationships are increasingly moving from operational to strategic in nature. The project represents an exciting direction for the development of the tools available to the growing Hydra Fedora community, facilitating the creation of a mediated publishing layer. More importantly, however, it has the potential to transform behaviors and relationships among important partners in the scholarly communication process:

- Through creating workflows that enable humanists to easily present rich electronic content either linked to their long-form publications or freestanding, the project opens the road to more innovative presentation of research by disciplinary communities who have often lagged behind their colleagues in science and engineering fields in engaging in digital publication.
- For university presses that are part of libraries, it creates a strong opportunity to extend those relationships beyond just administrative reporting or physical collocation to explore projects that truly leverage the complementarity between libraries and publishers.
- For libraries engaged in the creation of infrastructure for the management of data, the project provides rich case studies of how these initiatives can serve the needs of humanists as well as scientists and will encourage the framing of data repository development in terms of service provision rather than collection management.

The proposed project advances the sustainability of digital long-form publication in the humanities in a number of ways: From a business perspective, this project spreads the costs of the enriched forms of publication scholars are increasingly demanding across a broad range of institutional partners and leverages existing infrastructure development. It will develop a hosted solution to ensure longer term sustainability. While it most obviously engages with the challenge of enabling publications to interact on the web with primary sources and other related materials, the publication of the five case studies will also involve a rethinking of traditional production practices to allow granular linking to source materials, the involvement of rights experts in analyzing permissions needs, and analysis by data librarians of the level of preservation commitment that can be made to different types of linked material. Important issues of concern to publishers will also be worked through regarding the appropriate level, form, and documentation of peer-review for linked materials and the sales and marketing challenges of distributing both restricted and open access digital materials in a monographic environment designed for print products.

Bios:

Principal Investigators

Charles Watkinson, Director, University of Michigan Press / AUL for Publishing, University of Michigan Library. Charles joined University of Michigan in July 2014 after five years as Director
of Purdue University Press and Head of Scholarly Publishing Services in Purdue University Libraries. He was previously Director of Publications at the American School of Classical Studies at Athens (ASCSA). He has been PI or a lead investigator on several grants connected to the organization and dissemination of digital information received from the Mellon Foundation, National Endowment for the Humanities (NEH), and the Deutsche Forschungsgemeinschaft (DFG). His publications as an archaeologist include a presentation of linked interpretation and data from the Shala valley in Northern Albania published as part of the Linking Electronic Archives and Publications (LEAP) initiative.

Jeremy Morse, Director of Publishing Technology, Michigan Publishing. Jeremy joined University of Michigan in March 2005 as Electronic Publishing Programmer for the Scholarly Publishing Office, then assumed the head of the nascent Publishing Technology unit with the creation of the new Publishing division in 2010. He was previously Digital Technology Specialist for five years at Northwestern University, and also served four years in that institution's Course Reserves, where he ascended to department supervisor.

Gary Dunham, Director, Indiana University Press. Before arriving at Indiana University Press in October 2014, Gary served since 2010 as director of publications for the American Speech-Language-Hearing Association, where he transformed the 173,000-member association’s publications program. He was previously executive director of State University of New York Press from 2008 to 2010 and director of University of Nebraska Press from 2004 to 2007. He has a PhD in anthropology from the University of Virginia and has taught at the University of Nebraska, Creighton University and the University of Virginia.

Patrick Alexander, Director, The Pennsylvania State University Press. Patrick started as Director of Penn State Press in 2009 after two years as as associate director and editor-in-chief and as the co-director of Penn State University Libraries’ Office of Digital Scholarly Publishing. He has been involved in academic publishing for almost 30 years, having worked for Walter de Gruyter (Berlin/New York), Brill Academic Publishers (Leiden/Boston), and Hendrickson Publishers (Peabody, Massachusetts).

Jane Bunker, Director, Northwestern University Press. Before starting as Director of Northwestern in 2010 Jane was associate director and editor-in-chief of the State University of New York Press, where she held a number of positions since 1996. A graduate of St. Norbert College, Bunker earned a master's degree in philosophy from Fordham University. She is a Board member of the Association of American University Presses.

Doug Armato, Director, University of Minnesota Press. Doug has been Director at Minnesota since 1998. He was previously Associate Director and Book Division Manager at Johns Hopkins University Press for 10 years, Assistant Director and Marketing Manager at University of Georgia Press, and Assistant Marketing Manager at Louisiana State University Press.
Business Consultant

Brian O’Leary, Principal, Magellan Media Consulting Partners. Brian O’Leary works with a range of book, magazine and association publishers on issues related to content management, platform development and business planning. Before becoming a consultant in 1998, O’Leary served as senior vice-president and associate publisher of Hammond Inc., where he oversaw the enhancement and maintenance of the company’s digital database of the world. He came to Hammond after serving in a number of operational roles at Time Inc., where he oversaw production for weekly titles that included Time and Time International. Brian O’Leary holds an AB in chemistry from Harvard College and an MBA from Harvard Business School.

Library Technology Advisory Group

Robert H. McDonald, Associate Dean of Libraries, Indiana University
Jon Dunn, Director of Library Technologies, Indiana University
Carolyn Walters, Executive Director, Office of Scholarly Publishing, Indiana University
Maurice York, AUL for Information Technology, University of Michigan
Elaine Westbrooks, AUL for Research, University of Michigan
Kate McCready, Publishing Services Librarian, University of Minnesota
John Butler, AUL for Data and Technology, University of Minnesota
Susan Doerr, Operations and Business Development Manager, University of Minnesota Press
Julie Renee Rudder, Digital Initiatives Project Manager, Northwestern University Library
Patricia Hswe, Head, ScholarSphere User Services, Penn State University