

Omeka Commons: Preserving and Sharing Our Dispersed Digital Commonwealth

Overview

Large collecting institutions and their foundation backers have made good progress in recent years digitizing and making collections available online. Private corporations such as Thompson and Google have also done extensive work in this area. As impressive as these large-scale efforts are, materials found in large archives represent just a fraction of the unique archival materials in institutional holdings. Far greater—in numbers and even richness—are the combined collections found in small museums, libraries, archives, and historical societies around the world.

To fund each of these institutions individually to mount their collections online is not feasible. Even if it was financially feasible, it is not clear these dispersed online collections would be any more discoverable for users than they are in their physical form: they would still be dispersed across the internet as they are currently dispersed across the landscape. Yet a centralized effort to assemble these collections in a single digital archive would be equally, if not more expensive, and would trample local sensitivities about ownership of cultural heritage.

How then to uncover, aggregate, preserve, and make available for public use this dispersed and largely hidden “national archive?” How to build a permanent collections resource that respects local control and is largely locally funded, but manages to be centrally available and discoverable by all?

The answer is to provide a low-barrier, low-cost, standards-based means for local institutions and individuals to mount their own online collections and a similarly low-barrier means for them to voluntarily share those materials for central preservation and discovery.

The Center for History and New Media’s *Omeka* <omeka.org> web publishing platform has made great strides with regard to the first half of this equation. In the year since the launch of the first public beta of *Omeka*, the free and open source software has helped hundreds of cultural heritage institutions and individual scholars mount online archives and exhibitions. The second part of this equation—providing a point of permanent preservation and central discovery—is the next challenge. In partnership with the Internet Archive, the Center for History and New Media aims to meet this remaining challenge by networking the thousands of local websites using *Omeka* and aggregating their contents into a single, permanent, open access repository: *Omeka Commons*.

Specifically, *Omeka Commons* will:

- 1) give every current and future *Omeka* user an easy, one-click option for adding her collections to a central preservation grade repository, providing institutions with both a preservation pathway for their materials, and greater community exposure and engagement;
- 2) provide researchers with immeasurably greater and centralized discovery and open access to the small cultural heritage collections contained in *Omeka* sites across the web.

Meeting these aims will require significant work in three areas: enhancements to the core *Omeka* software, a new *Omeka Commons* server (to aggregate participant collections, serve researchers, and interface with the Internet Archive’s preservation repositories), and the design and production of the *Omeka Commons* discovery portal. The Center for History and New Media (CHNM) requests \$984,099 from IMLS to accomplish these important tasks, disseminate the new products, and evaluate the results.

I. Assessment of Need

Few small museums, libraries, archives, and historical societies share any collections online. During the past year *Omeka*—a free, open source, standards-based, collections-focused web publishing platform—has empowered hundreds of smaller cultural heritage institutions to achieve a more professional, more sustainable web strategy and online collections presence. Since its public

beta release, *Omeka* has been downloaded nearly 5000 times, and its users have cataloged tens of thousands of digital objects with standardized Dublin Core metadata. Thousands more are being added every month, and the growth of both users and objects is accelerating. Currently a server-installable download, *Omeka* will launch a “freemium” (free or nearly free) subscription web service in 2009 at *Omeka.net*. The launch of *Omeka.net* will make setting up an *Omeka* site as simple and affordable as setting up a Blogger or Flickr account. We anticipate the numbers of institutions, especially small institutions, using *Omeka* to rise exponentially once this web service becomes available. (See also Appendix B, About *Omeka*.)

The *Omeka* team has joined in many celebrations as smaller museums, libraries, archives, and historical societies have mounted online collections and exhibitions over the past several months. But as important as it is, simply publishing materials online is usually not the end of an institution’s web worries. How will they preserve their newly minted digital materials for the long term? How will researchers find and make use of the fruits of their hard labors in a World Wide Web with more than one trillion pages?¹ *Omeka Commons* aims to address these challenges. (See Appendix A for Endnotes.)

1) Museums, libraries, archives, and historical societies need open, low-cost, long-term digital preservation solutions

The issue of long-term preservation of digital records and assets is a tremendous problem for all cultural heritage organizations. Over the course of the past decade, this problem has been well studied and innovative solutions have started to emerge in the form of new standards (Dublin Core, OAI, JPEG2000), new communities of interest (CNI, DLF, MCN), new private/public collaborations (Google Books’ partnership with university research libraries, Footnote’s partnership with NARA), and new tools (Fedora, D-Space). However the great majority of the research, resources, and effort in the area of digital preservation has been focused on a relatively small number of very large institutions (research libraries, for example, or larger art museums). The nation’s thousands of smaller museums, libraries, archives, and historical societies remain far behind their larger cousins. A 2005 IMLS-funded study by the Northeast Document Conservation Center of digital preservation practices of mainly small and medium-sized museums showed that though more than 90 percent of institutions have digitized collections, only 13 percent have written plans or policies for preserving these collections. More than 60 percent of these institutions allocate no or very low levels of funding to managing and preserving digital collections. The study concludes that though “digital preservation is on their radar,” there is nevertheless “a general lack of understanding...of even the most rudimentary aspects of digital preservation” among these institutions. It is clear that smaller and medium size cultural heritage institutions have neither the financial resources nor the staff expertise to meet all of the problems they face in house.²

Some of the lessons learned and strategies pioneered by the large institutions among the digital preservation vanguard may translate to help smaller institutions make some steps at home. Flexible, lightweight metadata standards such as Dublin Core are a good example. Others clearly will not help: it is unlikely that any county historical society could configure and maintain its own installation of Fedora or D-Space, for instance. Indeed, even more modest aims—for example piggybacking on the D-Space implementation of a nearby research university—can be stymied by administrative hurdles and persistent “structural and terminological incompatibilit[es].”³

Others remain untested. For example, public/private partnerships hold out hope to some institutions as a way to outsource the financing and staffing of their digitization and digital preservation projects. Yet in these difficult economic times the potential that commercial entities might decide to walk away from ventures with nonprofit cultural institutions is high. The partners of the HathiTrust recently realized just this: that the digitization work they undertook as part of their Google Books partnerships would be put at risk should Google abandon the project or go bankrupt. Thus HathiTrust brought together all of the universities that contributed to the Google Books project and provided a way for them to aggregate and preserve more than two million digital copies that they received from Google. As John Wilkins, associate university librarian at the University of Michigan and executive director of HathiTrust explained, “This is a commitment to the permanence of the materials.” If small and medium sized cultural institutions are to outsource their digital preservation requirements, they likewise need a stable, not-for-profit commons to save and protect their resources.⁴

2) Scholars and researchers need improved discovery and open access to locally held collections

Millions of objects, photographs, and documents sit in the storage facilities of museums, libraries, archives, and historical societies, under-utilized by scholars and researchers. Digitization and the web have long promised to expose these valuable

collections to the light of day. Indeed *Omeka* is designed to aid smaller cultural institutions in this very effort.

But discovery remains a problem. The Heritage Health Index reports that it is impossible to gather even basic information about the condition of more than half of the digital materials in the U.S. As powerful as it is, Google does not do a good job of isolating collections material from items of more general interest. As Timothy Cole and Sarah Shreeves have noted in reference to IMLS's own *Digital Collections and Content* registry: "The benefit of having a rich diversity of quality and authoritative information available online is clear, but the magnitude of that benefit is tempered for many end users by the difficulties in locating specific, desired information resources within the almost overwhelming aggregation of information now available. Every week there is more useful information to find, but also every week, the amount of information that must be sorted through to find specific information desired grows as well."⁵

Cole and Shreeves' *Digital Content and Collections* project <<http://imlsdcc.grainger.uiuc.edu/>> is an ambitious and well-executed step towards helping alleviate what Roy Rosenzweig called this "problem of abundance." Yet the mandate of that project only extends to providing centralized search and discovery of digital collections funded by IMLS National Leadership Grants. The Flickr Commons partnership <flickr.com/commons>, another attempt to provide greater visibility to cultural heritage collections, is likewise limited to larger institutions such as the Boston Public Library and the Brooklyn Museum. Smaller institutions do not have access to similarly centralized discovery facilities.⁶

In her summary of the Web Wise 2007 conference on "Stewardship in the Digital Age" Diane Zorich writes, "the reality of information access today is that if we continue to isolate or 'silo' our collections on our web sites, we limit rather than enhance access." This is especially true for the siloed online collections of smaller institutions. *Omeka Commons* is designed to open those silos.⁷

II. National Impact and Intended Result

1) *Omeka Commons* will provide long-term digital preservation for cultural heritage institutions, raise visibility, and maintain local control over digital assets.

Omeka Commons will give every institution using *Omeka*, both now and in the future, a choice of adding its collections to a central preservation repository and an easy, one-click solution for doing so. The impact on participating institutions across the nation and around the world will be three fold:

First, *Omeka Commons* will provide these institutions with the long-term digital preservation solutions they cannot provide themselves but which even local funding agencies (such as state humanities councils) are increasingly demanding as a condition for funding digital projects. *Omeka*'s standards-based architecture and best-practice-encouraging user interface will insure that materials in dispersed archives are loaded into the system in a way conducive to archiving. A new REST API (Representational State Transfer Application Programming Interface) will make it easy to update the *Commons* as participants add materials to their *Omeka* websites.

Second, and more critically, *Omeka Commons* will allow participating institutions to maintain a strong measure of local control because their materials will be located first and foremost on their own *Omeka* websites. A copy will be preserved and served at *Omeka Commons*, but it will be done so with clear attribution and links back to the original site. A choice of open access Creative Commons licenses (and guidance for using them) at the point of contribution will help participating institutions assert their copyrights and clearly lay out terms of use for the items they have added to *Omeka Commons*. Moreover, even within the *Omeka Commons* universe, each institution will be provided with its own "channel" page containing institutional information, its logo, links, and a browseable catalog of the materials it has added to the *Omeka Commons*. Throughout this entire process, institutions will maintain complete control over which items they share with the *Commons*, the reproduction quality at which they are made available, and the way they may be used by the public. In this way, the centralized/decentralized approach of *Omeka Commons* will help solve small institutions' preservation and discovery problems centrally without compromising local control of cultural heritage. (See Use Cases, Appendix C.)

Finally, *Omeka Commons* will provide increased visibility to local collections and provide a means of building community around objects. Researchers who discover local materials through the *Commons* will be provided a direct link back to the original

source, driving traffic to local websites. Institutions will also have a place where researchers can tag and comment on items they find (similar to facilities larger institutions enjoy in the Flickr Commons project), driving crowd-sourced collections information back to the institution and building a community around their collections materials. Recent studies, such as those by the Library of Congress, the Powerhouse Museum, and the *steve.museum* project, confirm the importance of these online user interactions to cultural heritage institutions. *Omeka Commons* will offer smaller institutions a measure of exposure and audience interaction currently available only to larger and better-funded institutions.⁸

2) *Omeka Commons* will improve discovery for researchers, increase access to cultural heritage collections, and encourage collaboration.

Facilitating research is most often cited as the reason cultural heritage institutions undertake digitizing their collections, according to a survey of IMLS-funded digitization projects that identified scholars as their intended audience.⁹ Yet digitized collections are only useful to researchers if they can find, use, and share these materials with peers and colleagues. *Omeka Commons* will help on each of these fronts.

First, *Omeka Commons* will provide a central place for researchers to begin their work with digitized cultural heritage material. The user interface will allow them to easily search for materials and access not only the item record, but also the actual digital files. Once researchers have located the material that they are searching for, they will be able to syndicate the results of their search in their own sites or applications using an RSS feed or simple YouTube-like “embed” string. This will exponentially increase the number of individuals exposed to items from participating institutions’ collections, while always keeping those items linked to the originating museum, library, archive, or historical society.

Second, open access is critical to making the *Commons* successful for researchers. ProQuest revolutionized periodical searching by offering full-text article searches over twenty newspaper archives. This is also true of the major archival databases, such as Chadwick-Healey’s Archives Finder, which lists the holdings of major archives and special collections based on their finding aids (but does not provide actual access to the material.) But unless a researcher belongs to an academic institution or can visit a major public library that pays subscription fees, she will not easily be able to locate these invaluable sources. More importantly, these entry points do little for the scholar or enthusiast who is interested primarily in non-documentary holdings. *Omeka Commons* will provide open access from one portal to valuable material cultural, media, and archival resources across institutions and repositories.

Third, *Omeka Commons* will not only expose scholars to new material for their work, it will also facilitate increased collaboration among scholars. Through intuitive sharing and recommendation tools in the *Commons*, scholars will be able to communicate with one another about items, collections, and ongoing scholarly work. Additionally, the *Commons* will provide scholars with a stable URI for each item, dramatically improving their ability to find, share, and cite material. The usage statement that accompanies the Creative Commons license choice will also allow scholars to easily understand the terms of their usage of material. Institutions like CHNM have considerable experience building collaborative digital projects that develop new scholarship based on resources found in museums, libraries, and archives that were collected, presented, and interpreted online.¹⁰ *Omeka Commons* will give others the same tools, both to find and use resources currently stuck in “silos” and to find institutions and individuals to collaborate with on new digital or analog projects.

Scholars are not the only researchers who form an audience for digitized collections. As so many cultural heritage institutions know, there are thousands of truly committed enthusiasts in the general public who are eager to have better access to materials related to space exploration, early scientific instruments, or Jacob Lawrence’s art. Students and teachers make up another core audience for cultural heritage outreach. With increased access to collections, a teacher will be able to discover heretofore unavailable materials to include in rich classroom activities. *Omeka Commons* will provide enthusiasts, students, and teachers with unprecedented access to the collections they desire and will deepen their connection to the home institution by driving traffic back to local sites. Additionally, *Omeka Commons* has the potential to introduce these visitors to new cultural heritage institutions, some of which might turn out to be in their own backyards. (See Use Cases, Appendix C.)

As the Obama administration readied to take office and a new Congress struggled with the requirements of a trillion dollar recovery package, the Association of Research Libraries called on these leaders to establish a new “digital, universal library or digital data commons” that could “span the collections of research libraries, museums, and other cultural memory organizations”

and “open the doors of these institutions to the public without regard to where they live.” *Omeka Commons* will advance this important plank of the renewal agenda, spanning the collections of smaller institutions, opening their riches to exploration and discovery by all, and moving us one step closer to the goal set forth by the American Council of Learned Societies of placing “the world’s cultural heritage—its historical documentation, its literary and artistic achievements, its languages, beliefs, and practices—within the reach of every citizen.”¹¹

III. Project Design and Evaluation Plan

Design

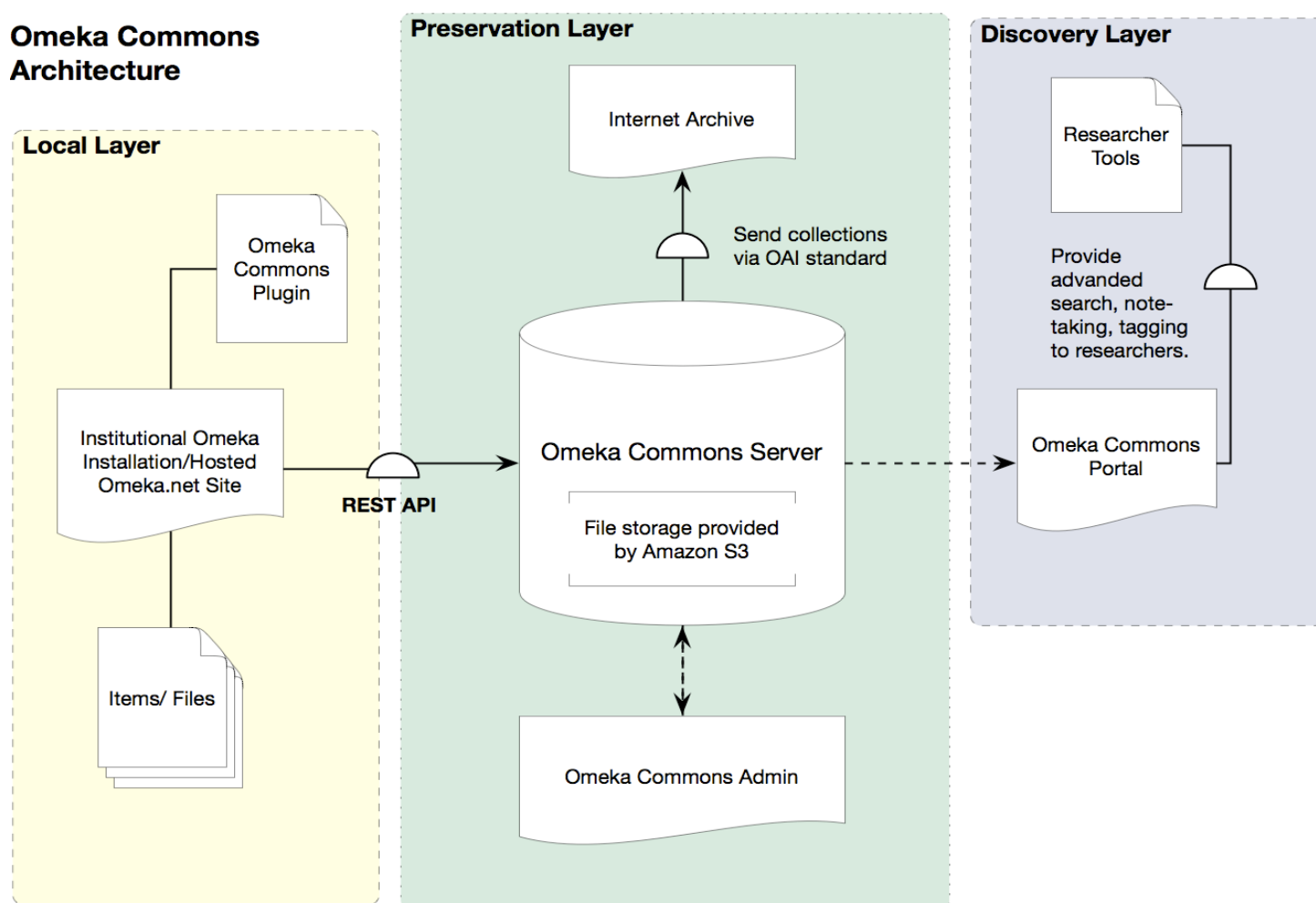
Meeting these aims will require significant work in three areas: enhancements to the *Omeka* software, a new *Omeka Commons* server (to aggregate participant collections, serve researchers, and interface with the Internet Archive’s preservation repositories), and the design and production of the *Omeka Commons* research portal. Throughout the project we will leverage extensive IMLS, Sloan, Mellon, and Kress-funded work on *Omeka*, but also the lessons learned and significant investments that have gone into the development of *Zotero Commons*—another collaboration with Internet Archive which, like *Omeka Commons*, seeks to uncover and preserve a “hidden” archive, though instead of the hidden archives of institutions, *Zotero Commons* takes on the personal file cabinets of the hundreds of thousands of individuals using CHNM’s *Zotero* <zotero.org> personal research management software. (See Appendix D for more information about *Zotero* and *Zotero Commons*.) The user interface elements, user authentication systems, data transfer protocols, preservation workflows and architectures, and researcher access tools for the two projects share much in common, and we will explore and exploit staffing and technology overlaps and efficiencies where beneficial to both projects. (*Zotero Commons* is funded by a major grant from the Andrew W. Mellon Foundation. See also, Budget Justification.)

Local Layer — Omeka software enhancements. The first work of the project will be making important additions to the *Omeka* software itself to accommodate the institutional participation, sharing, and data transfer that will be at the core of the *Omeka Commons* project. This will take the form of a new plugin that will enable individual *Omeka* websites to write collections data to *Omeka Commons*. Pre-installed with all *Omeka* and *Omeka.net* websites, this plugin will add an additional dialog to the standard *Omeka* setup process asking whether institutions want to participate in the *Omeka Commons* (See Appendix E, *Figure 1*). If the answer is “Yes,” the user will be directed to an application form that is sent to the *Omeka Commons* administrator for review and approval. If “No,” the plugin will deactivate, but will remain in the plugins directory should the user wish to reactivate the plugin and join *Omeka Commons* later. Once an application has been approved, institutions will receive an authentication key to paste into the plugin configuration panel allowing them to contribute to the *Commons*. A panel on the “edit items” form in the *Omeka* administrative interface will allow an institution to add an individual item to the *Commons* and to choose a Creative Commons license for that item (See mockup Appendix E, *Figure 2*). A panel on the “edit collections” form will allow an institution to add all items in a particular collection to the *Commons* and choose licensing terms for the entire collection (See mockup Appendix E, *Figure 3*).

Preservation Layer — Omeka Commons server. When local institutions choose to add material to the *Commons*, item level metadata and associated files will be written via a fully functional REST API to a central server, where they will be assigned unique identifiers and filenames to prevent overwrites and facilitate preservation. This server will be housed in a secure, enterprise-grade data facility (including full UPS power, generator backup, N+1 redundancy, and precision, high-density HVAC) directly atop the main internet backbone in Ashburn, VA. A robust production-grade MySQL database will store aggregated collections data, drawing on Amazon’s S3 cloud storage solution <<http://aws.amazon.com/s3/>> for file storage. Data on the *Omeka Commons* server will be managed through a custom administrative interface, where CHNM staff will approve and manage institutions and users, and view and organize contributed collections data (See mockup Appendix E, *Figure 4*). Meanwhile the server will provide a regular stream of collections data via OAI to the Internet Archive’s main Data Repository for preservation, where they will also be made available for discovery and use at the Internet Archive’s main website <www.archive.org>. Among other measures, Internet Archive will archive digital files, store multiple copies of collections materials, maintain preservation policies to secure the long-

term accessibility of *Omeka* materials, and perform periodic hardware and software migrations to insure long-term preservation of *Omeka Commons* collections. (For more on Internet Archive’s preservation architecture, see Appendix F.) To deploy and manage these systems and protocols, we will share significant expertise, server technology, infrastructure, and staff resources with the *Zotero Commons* collaboration.

Omeka Commons Architecture



Discovery Layer — Omeka Commons discovery portal. CHNM will create a feature rich discovery portal to enable researchers to find and use collections in the *Commons*. The home page <commons.omeka.org> will display a list and RSS feed of recent items, a tag cloud, a simple search, and a login panel (See mockup Appendix E, *Figure 5*). From the home page, links will allow researchers access to the entire collection through a faceted browse interface (including originating institution, item type, collection, subject, and Dublin Core metafields), and an advanced iTunes-like query builder. From the search results, researchers will access individual item pages, complete with exposed DCMI and media files (See mockup Appendix E, *Figures 6, 7, 8*). Results will be fully readable by *Zotero*, which will give researchers access to their results through the *Zotero* Firefox add-on and multi-platform web service. They will also be available for subscription or syndication via RSS and YouTube-style “embed” strings. Researchers who sign up for an *Omeka Commons* user account will enjoy added social features: under a tab entitled “My Commons,” researchers will be able to save items to a personal collection, annotate items, tag items, and share selected items with colleagues and friends (See mockup Appendix E, *Figure 9*). They will also be able to participate in public discussions of items similar to discussions on Flickr Commons. In turn, institutions contributing collections to the *Commons* will be able to engage users of the discovery portal through customizable institutional “channels”—individual homepages that will include an institution’s description, logo, chosen color palate or background image, featured items, institution-specific search, and institution-specific RSS feed (See mockup Appendix E, *Figure 10*).

Work Plan

Omeka Commons is an ambitious proposal, but one that can be accomplished in three years in five overlapping phases:

Phase I (9 months, December 2009-August 2010) — Local layer, Omeka software enhancements. We will begin the first phase of work on the project by assembling a detailed requirements checklist and development timetable for enhancements to the *Omeka* software, reaching out to *Omeka*'s active open source developer community for formative assessment of our plans. We will then begin development and testing of the *Omeka Commons* plugin, which will include an application to participate, programming and interface elements for adding items and collections to *Omeka Commons*, and Creative Commons licensing options.

Phase II (9 months, June 2010-February 2011) — Preservation layer, Omeka Commons server design and construction. The second phase of the project will start with construction and implementation of a detailed data model for the *Commons* server. Next, we will build the REST API and authentication system necessary to enable data transfer between local *Omeka* and *Omeka.net* sites and the *Omeka Commons* server. When this work is completed, we will implement the Amazon S3 storage solution, create and test an OAI bridge to the Internet Archive preservation repositories, and design the administrative interface for the *Commons* server.

Phase III (9 months, December 2010-August 2011) — Discovery layer, Omeka Commons portal design and construction. We will begin the third phase by engaging a series of focus groups to determine what researchers will need from the *Omeka Commons* discovery portal. Using recommendations from focus groups, we will compose wireframes and mockups to guide development. Then, we will implement basic browse and display functionalities, advanced search facilities, RSS and embed syndication, "My Commons" features, and institutional channel pages. Prior to launch, we will design and conduct a series of user testing routines.

Phase IV (9 months, June 2011-February 2012) — Beta launch, user feedback, revision. We will rely on *Omeka*'s robust developer and end user community to help us test the *Commons beta*. Additionally, we will invite individuals from the museum and academic communities to become user-testers. In order to insure the integrity of preservation data and prevent data loss, we will conduct regular checks of our data transfer protocols (both from local *Omeka* sites and to Internet Archive) and preservation architecture. Throughout this period we will continue to make iterative improvements and enhancements to the local, preservation, and discovery layers of the *Omeka Commons* technology to insure maximum uptime, a seamless user experience for both local institutions and researchers, and long-term preservation.

Phase V (1 year, December 2011-November 2012) — Dissemination and ongoing evaluation. We will leverage our extensive contacts within the cultural heritage community to locate participants and to encourage use of the *Omeka Commons*. In addition, we will publicize the project through a variety of print media and conference presentations. (See also, Section V, Dissemination, below.) Finally, during this period and afterward, we will deploy a range of formal and informal evaluation measures and continue to perform iterative upgrades to the *Omeka Commons* technology. (See also, Evaluation, below.)

Evaluation

CHNM has a long track record of iterative evaluation in software and web development projects, gathering, responding to, and implementing feedback from users, community developers, and other stakeholders at each stage of planning, development, and dissemination. Evaluation of the *Omeka Commons* project will follow this model, and fall into three main categories:

First, we will engage in an agile development process that will allow for many points of testing, feedback, and mid-course corrections during Phases I, II, III of the project. As we make changes to the core of the *Omeka* software, we will run frequent automated unit testing. These results will assist the developers in producing robust code. We will also conduct two rounds of live user testing that will inform UI changes to *Omeka* and construction of the discovery portal. We will solicit volunteers to participate in more informal testing from *Omeka*'s active open source developer community, the cohort of institutions already using *Omeka*, and the community of scholarly researchers. To insure long-term collections integrity, we will perform frequent systematic checks of our data transfer protocols and preservation architecture.

Second, we will conduct two feedback surveys with cultural heritage institutions during the *Omeka Commons* beta launch (Phases IV). The first survey, conducted during Fall 2011, will provide feedback on the upload and preservation process. The second survey, conducted during Spring 2012, will focus primarily on the *Omeka Commons* portal, discoverability, and researcher

interface features. The results of these two surveys will assist the web development team in the final round of revisions to the *Omeka Commons* system as a whole.

Third, we will maintain open lines of communication with all of our users (Phases IV-V). Just as we have in the course of the *Omeka* project, we will maintain Google groups, a project blog, consistent IRC chat hours, and open documentation. Additionally, there will be a prominent feedback feature in the *Omeka Commons* portal design so that users can easily share their thoughts and reactions with the development team.

IV. Project Resources: Budget, Personnel, and Management

Budget

The total cost of the *Omeka Commons* project is \$1,968,266. Of this, \$984,167 or 50% is non-federal cost share, which includes contributions of staff time and infrastructure from CHNM and Internet Archive. We are requesting the remaining amount of \$984,099 from IMLS. This contribution will pay for project staff at CHNM and Internet Archive, research assistants, technology services, travel, evaluation and dissemination costs, and mandated indirect costs. All IMLS funds will be managed by CHNM and GMU, which have a long track record of managing federal grant money and doing cost efficient, high quality web and software development work.

Personnel

Tom Scheinfeldt (Executive Producer) has shared responsibility for the program vision and particular oversight of internal staffing, technology development, quality assurance, budget and grant management, and long-term sustainability.

Sharon Leon (Executive Producer) has shared responsibility for the program vision and particular oversight of institution relations, end user development and testing, program evaluation, and product dissemination.

Dan Cohen (Producer) will provide overall guidance and consultation to the project based on his experience in developing *Zotero* and *Zotero Commons*.

Sheila Brennan (Project Manager, Outreach) will coordinate partner meetings and involvement, museum outreach, product dissemination, and evaluation.

Jeremy Boggs (Project Manager, Development) will coordinate all programming and web development activities and provide specific expertise on user interface design for *Omeka* enhancements and the *Omeka Commons* discovery portal.

Kris Kelly (Senior Programmer) will implement *Omeka* code enhancements (including primary development of the *Commons* plugin) and oversee programming efforts on the *Omeka Commons* server and discovery portal (including primary development of the REST API).

Jim Safley (Digital Archivist/Metadata Specialist) will oversee the design and creation of the data model for the *Omeka Commons* server.

Dan Stillman (Senior Programmer) will design and implement systems architecture for the *Omeka Commons* server (leveraging related work on *Zotero Commons*) and provide programming support and expertise in developing the *Omeka Commons* server and discovery portal.

Raymond Yee (Data Integration/API Specialist) will design and implement data transfer models between the *Omeka Commons* server and Internet Archive (leveraging work on *Zotero Commons*) and provide advice and programming for the *Omeka Commons* REST API.

Linda Frueh (Project Manager, Internet Archive) will coordinate program activities at Internet Archive, including collections and data repository work and staff, budgeting, dissemination efforts, and partner relations with CHNM.

Management Plan

CHNM's extensive experience in digital projects equips it to manage the planning, development, and dissemination of *Omeka Commons*. The project builds directly on three core strengths of CHNM: 1) our development of widely used digital tools, including *Omeka* (recent winner of a Mellon Award for Technology Collaboration) and *Zotero* (named among PC Magazine's

“Best Free Software” two years running); 2) our creation of large-scale, preservation-grade digital archives, such as the *September 11 Digital Archive*, which have fostered an interactive model for documenting and preserving the past; 3) our use of the web for more than a dozen years as a vehicle for engaging broad and diverse audiences, often in partnership with museums and libraries, for example on our *Gulag: Many Days, Many Lives* project (a partnership with the National Park Service) or on the IMLS-funded *Object of History* project (a partnership with the National Museum of American History). Internet Archive likewise brings a more than decade-long track record of successful project management expertise to *Omeka Commons*.

CHNM’s success results from combining the knowledge and skills of credentialed historians with the experience of our multimedia and technical staff and considerable technical and equipment resources. *Omeka Commons*’ executive producers, Tom Scheinfeldt and Sharon Leon have overseen work on numerous grant projects, including the *Omeka* project, with budgets totaling \$20 million. Sheila Brennan has had considerable hands-on experience managing digital collecting projects and public history websites. Jeremy Boggs is well established as leader in the field of digital humanities design and development with work on dozens of websites and scholarly software development projects. In addition, the project will benefit from the extensive expertise of Dan Cohen, Director of CHNM and Linda Frueh, Regional Director of Internet Archive, whose experience on the *Zotero Commons* project will provide valuable management templates to the *Omeka Commons* project.

Day-to-day work of the project will be managed by way of weekly production team meetings, periodic consultations with partners, and frequent communication via email, blog, instant message, and conference call. Open communication among the producers, managers, and staff will ensure that we integrate the comments of partners into production, and that we meet all of our project deadlines.

V. Dissemination

CHNM’s ability to reach institutional and researcher audiences is proven by its long track record of delivering tools and resources to end users. CHNM’s websites receive more than 11 million unique visitors each year, and more than one million cultural heritage professionals and scholars use CHNM’s tools. We will build on this strong track record in the *Omeka Commons* project, harnessing existing audience connections and building new ones. Dissemination of information about *Omeka Commons* will reach both institutional audiences seeking to adopt *Omeka* and participate in the *Commons*, and researchers and scholars seeking open access to the resources available through the *Commons* portal.

To drive researcher traffic to the *Omeka Commons* and to encourage institutional participation, we will post announcements to museum listservs (e.g. MUSEUM-L), H-Net lists, and make use of the *History News Network*, a CHNM affiliate that sends a newsletter to 13,000 people three times per week. Equally important, CHNM sponsors a number of widely read professional blogs, and these blogs have proven to be a very effective means for rapidly disseminating news to those interested in digital tools. We will highlight newly added *Commons* items periodically on the *Omeka* project blog. And, we will continue to promote the *Commons* by designing and passing out t-shirts, thumb drives, stickers, or other promotional items to reward *Commons* adopters.

We will also make use of the vast opportunities for electronic social networking available to us. In addition to the *Omeka* project’s existing channels—active forums, a frequently updated open documentation wiki, a well used dev list—we will establish and monitor a Google group for cultural heritage institutions and one for researchers and scholars to discuss their participation in and use of the project. Additionally, members of the development and outreach team will hold regular IRC chat hours to enable real-time response to user questions and concerns. Finally, we will heavily use the services of del.icio.us, Flickr, and Twitter to create a sense of involvement and community for our constituencies.

In addition to electronic outreach, we will reach out directly to museums, libraries, archives, and historical societies, something that is already a key component of our current strategy for promoting *Omeka*. Building on of strong relationships with institutions such as the Minnesota Historical Society, Ohio Council for the Humanities, and Smithsonian Affiliations, we will seek their assistance in promoting not only the adoption of the project by their cultural heritage networks, but also to point interested researchers to the *Commons* portal. CHNM will make presentations and organize panels about the *Omeka Commons* at major conferences, including Museums and the Web, American Association of Museums, American Association for State and Local History, American Historical Association, and Coalition for Networked Information. Finally, we will also schedule joint workshops, together with CHNM’s *Zotero* team, to teach researchers and scholars how to best use the *Omeka Commons* for

finding, saving, and sharing primary source materials.

We will also harness the enormous reach of the Internet Archive to disseminate *Omeka Commons*. The Internet Archive receives more than 300 page requests *per second* for its Web Collection alone—over 25 million requests per day—and is serving its digital content from all collections at the average rate of 6 Gigabits per second. The Open Content Alliance (OCA), a collaborative organization founded by the Internet Archive and dedicated to the creation of joint online collections will be a particularly strong conduit for outreach. The OCA currently has 110 contributor institutions—cultural, academic and corporate—that share the values of open sharing of content, and it is continuing to grow. The Internet Archive and its partners in the OCA maintain a listserv, a website <opencontentalliance.org> and host an annual meeting to share progress on new initiatives. These mechanisms will all be used to share progress on the *Omeka Commons* project and engage community input.

VI. Sustainability

CHNM's ability to sustain the *Omeka Commons* is demonstrated by more than a dozen years of work in the field of digital history. Since 1994, CHNM has used digital media and computer technology to present and preserve the past, and it currently hosts more than four dozen online history projects, among them several large, highly regarded and highly visited digital archives and preservation projects such as the *September 11 Digital Archive*, the *Hurricane Digital Memory Bank*, and the *Papers of the War Department, 1784-1800*. Particularly telling of CHNM's stability is that the creators of some major digital history projects—such as the award winning *DoHistory*—have given their projects to CHNM in order to insure their permanence.

In recent years, CHNM has also become a major developer of digital tools in the humanities with support from IMLS and the Sloan and Mellon foundations. Most notably, CHNM has developed *Zotero*, an innovative, open source reference management and note taking tool, and *Omeka*, which forms the foundation for the *Omeka Commons* proposal. The open architecture and standards of *Omeka*, its open source license, its core open source technologies (MySQL, PHP), and especially its strong and growing open source community underwrite *Omeka's* future development and long-term sustainability. Less than one year after its public beta release, *Omeka* has already engaged a large and active developer and user community, something recognized in December 2008 when *Omeka* was awarded a \$50,000 Mellon Award for Technology Collaboration, which will be allocated toward endowing year around user support for the project. While committed to keeping *Omeka* free and open, CHNM is also exploring several cost recovery models for the project, in particular a “freemium” subscription model for the *Omeka.net* hosted web service. (See also Appendix B, About *Omeka*.)

CHNM's overall financial stability stems from several sources. With the help of two NEH Challenge Grants, CHNM has raised a nearly \$3 million endowment, which provides for CHNM's long-term viability as a unit of the George Mason University's Department of History and Art History. Mason also has a strong institutional commitment to CHNM, which has a \$2 million annual budget and a staff of more than fifty. The History Department has made “new media” a central focus of its new PhD program, which insures a steady stream of students (at least six of whom work as Graduate Research Assistants at CHNM) interested in sustaining existing projects and creating new ones. This substantial endowment, demonstrated fundraising success, strong institutional support, and solid technology allows CHNM to guarantee that the *Omeka Commons* will be a permanent and reliable resource for cultural heritage institutions and scholarly researchers.

The long-term sustainability of the Internet Archive has been demonstrated by twelve years of successful operation as a trusted, nonprofit digital repository of record with a broad funding base. Underwriting this stability is a successful program of fee-for-service web harvesting, book scanning and microfilm digitization, and content hosting as well as a track record of securing both federal and private foundation grants. The Internet Archive will store, maintain, migrate, and preserve fully functional digital files and metadata for *Omeka Commons* collections in perpetuity. Since the *Omeka Commons* data and files will be stored in the Internet Archive's main Data Repository, in the same format and using the same clustered storage environment and management software as its extensive existing collections, these files will automatically be handled with the same preservation and migration policies of all the Internet Archive collections. The value of storing these items at the Archive is that they will, as long as the Internet Archive's collections exist, be maintained with integrity checks and long-term preservation activities.

(For more information on the Internet Archive's long-term preservation architecture and protocols, see Appendix F, Preservation Architecture. See also, Organizational Profiles.)