Omeka S, Enhanced Description and Dissemination
Abstract

The Roy Rosenzweig Center for History and New Media (RRCHNM) at George Mason University is seeking support from the Institute of Museum and Library Services through a National Leadership Grant for Libraries to embark on "*Omeka S*, Enhanced Description and Dissemination." This project builds upon the strong foundation provided by *Omeka S*, RRCHNM's linked data web-publishing platform for medium and larger cultural heritage institutions, to take the next step in supporting the networked, semantic description, and extensive sharing of digital cultural heritage collections, by completing the following deliverables between October 2015 and September 2018:

1. Developing an *Omeka S* module to integrate Library of Congress LOD vocabularies and authority files into the resource creation process
2. Developing an *Omeka S* module to integrate Getty Research Institute LOD vocabularies and authority files into the resource creation process
3. Developing an *Omeka S* module that allows administrators to create their own local controlled authorities to use in the resource creation process
4. Developing a Basic Description Resource Template for *Omeka S* that fulfills the requirements of the DPLA Metadata Application Profile and the Europeana Data Model.
5. Developing an *Omeka S* module that allows end-users to share individual resources with embed codes, email, or popular social media platforms
6. Conducting four workshops for GLAM developers to train them to customize and extend *Omeka S*
7. Supporting the cultural heritage community as they integrate *Omeka S* and the above proposed description and dissemination tools into their digital strategies

The work funded by this grant will result in five key results for the cultural heritage community that support the effort to build a national digital platform:

1. **Increasing the integration of Linked Open Data (LOD) authority files in metadata description for digital collections:** LOD vocabularies modules will help users create description that capitalize on the LOD universe through the use of controlled authority files from the Library of Congress and the Getty Research Institute. The use of these standardized description values will increase the discoverability of the digital collections being described by linking them in to the rich context of the growing semantic web.
2. **Increasing the ability of cultural heritage organizations to implement their own local controlled authorities:** The Local Controlled Authority module will help GLAMs undertake the process of standardizing their description, and contribute to the development of authorities that might eventually become standards in the field.
3. **Increasing the likelihood that newly created metadata for digital collections can be smoothly transferred to key aggregators that make up the national digital platform:** By incorporating the minimum requirements for the DPLA MAP v.4.0 and the Europeana Data Model, the Basic Resource Description Template will help cultural heritage organizations begin the process of creating description in a way that supports easy data aggregation with fewer crosswalks and data transformations.
4. **Increasing the ease of circulation for digital cultural heritage collections through popular web and social media platforms:** The Sharing and Social Media module will allow GLAMs to capitalize on the willingness of their publics to boost the circulation and exposure of their digital collections by providing an easy way for visitors to *Omeka S* sites to push individual digital resources out to social media platforms, blogs, and digital publications.
5. **Increasing technical capacity of GLAM staff:** Through a series of four workshops, the *Omeka* team will work to train 100 interested GLAM developers to customize and extend the *Omeka S* platform to meet the needs of their individual organizations. This work will grow the community of professional who have moderate technical skills that can be harnessed in the service of making collections and collections data more publicly accessible.

Omeka S, Enhanced Description and Dissemination

1. Statement of Need

Since its initial description in 2006 as a web of data designed to facilitate use, reuse, and sharing, the semantic web has captured the imaginations of cultural heritage developers who dream of embedding their collections in an expansive contextual network of data. This vision of linked cultural heritage data, included in the *2010 Horizon Report* for museums as achievable within four to five years, would move resources and objects out of brittle, siloed repositories and make it possible for humans and machines to freely access available related materials. Five years after that prediction, we see that the *Horizon Repor*t for libraries (2014) projects that libraries are still four to five years away from wide adoption of semantic web technologies.[1] It seems as if the actual implementation of the Linked Open Data (LOD) to fulfill this vision may be perpetually delayed, even amongst medium and larger galleries, libraries, archives, and museums (GLAMs) with able technical staff.

There are a number of reasons for the creeping adoption of LOD within the cultural heritage fields. One clear reason is that libraries, archives, and museums have built up twenty years of digital artifacts and infrastructure that do not conform to the principles laid out by Tim Berners-Lee to guide the creation of stable, persistent LOD through Uniform Resource Indicators (URIs) and the Resource Description Framework (RDF).[2] A second reason, aptly pointed out by Dorothea Salo and Ed Summers in their 2013 paper examining the adoption of LOD by libraries, archives, and museums, is that the unique collections held by museums and archives have led those organizations to share only certain elements of standardized description, such as authority files and thesauri, rather than full records which is common in the library community.[3] Nonetheless, the attractions of LOD for cultural heritage organizations are clear: it has the power to stabilize access to digital collections and to offer new routes of access to those resources by contextualizing and embedding them in the ever-expanding web of data.

In 2012, the Roy Rosenzweig Center for History and New Media (RRCHNM) <http://chnm.gmu.edu> at George Mason University (GMU) started developing a new version of our popular open-source web-publishing platform, *Omeka* (*Classic)* <http://omeka.org/>, that was designed specifically with the principles of LOD in mind (*See Appendix A, Omeka Project History*). Developed for medium and larger GLAMs, *Omeka S* uses JavaScript Object Notation-Linked Data (JSON-LD) as its native data format, which makes it possible to enmesh *Omeka S* in the LOD world. Every *Omeka S* Resource (item, item set, media) has a URI, and the core software includes the following Resource Description Framework (RDF) vocabularies, which maximizes its data interoperability with other data publishers: Dublin Core Metadata Initiative (DCMI) Terms; DCMI Type; The Bibliographic Ontology (Bibo); and The Friend of A Friend Vocabulary (FOAF).[4] In addition to including these popular LOD vocabularies, *Omeka S* offer the ability to use the URIs for other *Omeka S* Resources as descriptive values within metadata fields, in essence linking one *Omeka S* Resource to another (i.e. using a Person type Resource for Thomas Jefferson as the value for the Creator field in description of the "Declaration of Independence" text Resource). Alternatively, a user

---

[1] Johnson, L., Witchey, H., Smith, R., Levine, A., and Haywood, K., *The 2010 Horizon Report: Museum Edition* (Austin, Texas:
[2] 2006: http://www.w3.org/DesignIssues/LinkedData.html
[3] Salo and Summers, "Linking Things on the Web: A Pragmatic Examination of Linked Data for Libraries, Archives and Museums," 2013: http://arxiv.org/abs/1302.4591, p. 4
[4] DCMI Terms: http://dublincore.org/documents/dcmi-terms/; DCMI Type: http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/; Bibo: http://lov.okfn.org/dataset/lov/details/vocabulary_bibo.html; and FOAF: http://xmlns.com/foaf/spec/.

could input a URI for an external resource (i.e. Thomas Jefferson's DBpedia page[5]). *Omeka S* also makes it possible for users to attach media to Resources in three ways: through a simple file upload, through the use of an embed code from an outside resource, or use of a URI for an existing resource. In our progress toward ensuring and maintaining interrelationships between humanities data, connectors between *Omeka S* and other systems will ensure that LOD is maintained. For example, the connector to Fedora fosters import of both Fedora's own vocabulary, and the W3C Recommendation Linked Open Data Vocabulary that it rests on. Together, these features prepare data in *Omeka S* to be fully embedded in the semantic web.

In February 2014, the Getty Research Institute (GRI) took a step forward to encourage cultural heritage organizations to embrace the promise of the semantic web by beginning to release their Vocabularies as LOD. President and CEO of the J. Paul Getty Trust, James Cuno pinpointed the potential gains of offering these descriptive databases as open content, "The release of the Vocabularies as Linked Open Data will help link cultural heritage information across cultures and between disciplines, leading to new advances in the humanities."[6] Currently, Getty offers the Art and Architecture Thesaurus (AAT), the Thesaurus of Geographic Names (TGN), and the Union List of Artists Names (ULAN) as SPARQL endpoints. With this venture, the Getty joined the Library of Congress in leading the field by offering free versions of their controlled authority files and standardized vocabularies. Since 2009, the Library of Congress (LC) has offered LC Subject Headings, LC Name Authority File, LC Classification, LC Genre/Form Terms, Thesaurus for Graphic Materials, AFS Ethnographic Thesaurus, Cultural Heritage Organizations, and several other files through their Linked Data Services portal.[7]

Yet, the promise that Cuno trumpeted can only be realized if individuals creating resource description have clear and easy ways to integrate the vocabularies into their workflows. While *Omeka S* helps to encourage the use of RDF vocabularies for resource description, it does not restrict the values that users input for description. Thus, the next step is to provide a way for administrators to incorporate standardized authorities into their workflows. LC and GRI have facilitated this work by publishing their established authority files as LOD. The work that remains is to bring the pieces together in the *Omeka S* resource creation/description interface.

Even with the existence of this large collection of widely used controlled authorities, these only represent vocabularies relevant for specific types of collections. Individual GLAMs are left crafting descriptions that are particular to their local contexts and collections. Yet even these values need to be standardized for systems where large numbers of individuals are working together through the process of description. Offering cultural heritage organizations easier ways to standardize description within their individual domains would provide the workspace for developing standards to play out in the field. A prime example of this need for standardization arises in the field of rights statements, where complex statements with uncontrolled values are inhibiting sharing and aggregation of digital resources, with the Digital Public Library of America (DPLA) citing over 85,000 different rights statements on their contributions alone.[8] While the answers to narrowing the ways that we communicate rights for digital

---

[5] http://dbpedia.org/page/Thomas_Jefferson

[6] James Cuno, "Art & Architecture Thesaurus Now Available as Linked Open Data," *Iris* (February 20, 2104), http://blogs.getty.edu/iris/art-architecture-thesaurus-now-available-as-linked-open-data/

[7] Getty Institute Vocabularies: https://www.getty.edu/research/tools/vocabularies/lod/index.html; LOC Linked Data Services Portal: http://id.loc.gov/; Kevin Ford, *Linked Data at the Library of Congress* , Presentation delivered at the SACO-At-Large meeting, American Library Association Annual Conference, Washington, DC (2010 June 27) http://id.loc.gov/static/presentations/kford-id-saco-ala-2010.pdf.

[8] Europeana DPLA Rights Statement Whitepaper: https://docs.google.com/document/d/1x10JsIfi8Y74pgJJEAqMtyO5iYp0p6DO5DrOZK-5umY/edit?usp=sharing

materials are still in flux, having an easy way for cultural heritage organizations to select a core set of statements and apply them in the course of publishing their resources takes a step in the right direction, and provides a stop-gap measure that can bring some level of coherence while the field waits for the establishment of an agreed upon set of statements.

While not everything about cultural heritage description can be standardized with controlled vocabularies, as large content aggregators such as DPLA gain traction in the cultural heritage landscape, institutions will require simple ways to shape their descriptive workflows to increase the likelihood that their digital collections can easily be shared with these entities. At present, national aggregators pull together metadata in an astounding array of schemas and states of completion. Offering templates and guides that conform to the metadata requirements common to these pillars of the national digital platform represents one way to transform the field by setting the default conditions for easier data aggregation. *Omeka S* provides a flexible resource templating system that allows administrators to select from the provided RDF vocabularies (or any that they have installed), but the platform does not yet offer a core resource template that supports the requirements that larger content aggregators have specified.

Of course, GLAMs are not only concerned with sharing their digital content through larger systems and platforms. They are also concerned with small-scale connections made between an institution and its online visitors, and encouraging personal interactions with their digital collections. Given the ubiquity of social media platforms and their role in the circulation of cultural heritage resources, there is an ongoing need for individuals to easily share a GLAM's digital cultural heritage resources, through popular platforms such as WordPress, Tumblr, Instagram, Facebook, Twitter, and Pinterest.

Finally, the development of software and tools to improve the description and dissemination of digital cultural heritage collections is only as good as the ability of GLAMs to implement and work with these solutions. Both cultural heritage technical and non-technical staff would benefit from more training specifically tailored to increase their ability to work with these platforms as they endeavor to share their digital collections with the widest possible public.

2. Impact

Given these description and dissemination needs, RRCHNM is requesting an $249,336 IMLS National Leadership Grant for Libraries under the National Digital Platform priority to support "*Omeka S*, Enhanced Description and Dissemination." The project aims to build upon the strong foundation provided by *Omeka S* to take the next step in supporting the networked, semantic description, and extensive sharing of digital cultural heritage collections by completing the following deliverables:

1. Developing an *Omeka S* module to integrate LC LOD vocabularies and authority files into the resource creation process
2. Developing an *Omeka S* module to integrate GRI LOD vocabularies and authority files into the resource creation process
3. Developing an *Omeka S* module that allows administrators to create their own local controlled authorities to use in the resource creation process

4. Developing a Basic Description Resource Template for *Omeka S* that fulfills the requirements of the DPLA Metadata Application Profile (MAP) and the Europeana Data Model[9]
5. Developing an *Omeka S* module that allows end-users to share individual resources with embed codes, email, or popular social media platforms
6. Conducting four workshops for GLAM developers to train them to customize and extend *Omeka S*
7. Supporting the cultural heritage community as they integrate *Omeka S* and the above proposed description and dissemination tools into their digital strategies

The work funded by this grant will result in five key outcomes for the cultural heritage community that support the effort to build a national digital platform:

1. **Increasing the integration of LOD authority files in metadata description for digital collections:** The set of LOD vocabularies modules will help users create description that capitalize on the LOD universe through the use of controlled authority files from LC and GRI. The use of these standardized description values will increase the discoverability of the digital collections being described by linking them in to the rich context of the growing semantic web.
2. **Increasing the ability of cultural heritage organizations to implement their own local controlled authorities:** The Local Controlled Authority module will help GLAMs undertake the process of standardizing their description, and contribute to the development of authorities that might eventually become standards in the field.
3. **Increasing the likelihood that newly created metadata for digital collections can be smoothly transferred to key aggregators that make up the national digital platform:** By incorporating the minimum requirements for the DPLA MAP v.4.0 and the Europeana Data Model, the Basic Resource Description Template will help cultural heritage organizations begin the process of creating description in a way that supports easy data aggregation with fewer crosswalks and data transformations.
4. **Increasing the ease of circulation for digital cultural heritage collections through popular web and social media platforms:** Leveraging the ubiquity of social media, the Sharing and Social Media module will allow GLAMs to capitalize on the willingness of their publics to boost the circulation and exposure of their digital collections by providing an easy way for visitors to *Omeka S* sites to push individual digital resources out to social media platforms, blogs, and digital publications.
5. **Increasing technical capacity of GLAM staff:** Through a series of four workshops, the *Omeka* team will work to train 100 interested GLAM developers to customize and extend the *Omeka S* platform to meet the needs of their individual organizations. This work will grow the community of professionals who have moderate technical skills that can be harnessed in the service of making collections and collections data more publicly accessible.

Over the course of the project, we will track and measure these outcomes in a number of ways. First, a prime indicator of a successful software development project is use. While we know from previous experience that it takes institutions some significant time to decide to adopt a new system or new features of an existing system, we

[9] DPLA MAP v.4: http://dp.la/info/developers/map/; Europeana Data Model: http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation.

expect to see a steady increase in *Omeka S* module use from the point of release onward. We will track the number of downloads of each module. Second, to measure the effectiveness of the modules to meet the needs of our key user base, we will implement a strong outreach strategy, and involve our active development community in reviewing and commenting on the code. We will also keep track of the ongoing discussions on the *Omeka* forums and the *Omeka* Dev list about the release candidates, and beta releases, integrating that feedback into the revisions of the modules for final release.[10] Finally, during the development, testing, and release cycles, we will ask for community feedback through targeted surveys that address user experience and functionality. Together, these efforts will help ensure that we achieve the goals stated above by creating modules that serve the needs of the community and that are easy to use.

We will employ a variety of metrics to evaluate the success of the four training workshops. First, we will ask participants to sign-up for the workshops prior to the events, which will offer us an opportunity to gather some demographic data and information about their work in digital cultural heritage. Second, at the conclusion of each workshop, we will ask participants to evaluate the training materials and their experience through a short online survey. This will provide us with vital data to revise and improve the materials and agenda between workshops. Also, we will encourage participants to make use of the established *Omeka* support channels for assistance after the workshops. Finally, we will survey participants several months after the workshop to get a sense of the ways they may have employed their new skills within their organization and their professional networks. Together, these efforts will allow us to refine our approach and measure the impact of our training efforts by the close of the grant period.

3. Project Design

When RRCHNM began working on the design for *Omeka S* in 2012, the development team underwent a thoroughgoing process of rethinking the basic *Omeka* software architecture. This required a willingness to leave behind some of the core infrastructural elements of the existing stand-alone version of *Omeka*, but the benefits of this open rethinking have been substantial. Built using PHP 5.4 with Zend Framework 2, Doctrine, and Composer, *Omeka S*'s Read/Write REST API enables all of the major software actions, such as the creation of users, items, sets, and sites. *Omeka S* uses JavaScript Object Notation-Linked Data (JSON-LD) as its native data format and ships with popular RDF vocabularies (DCMI Terms, DCMI Type, Bibo, and FOAF), which makes it possible to enmesh *Omeka S* in the LOD world. *Omeka S* is also insulated from cross-site scripting and cross-site request forgery vulnerabilities, bolstering its security for deployment in an institutional setting. Additionally, modules facilitate the flow of materials from Fedora and DSpace repositories, and Zotero collections into the *Omeka S* environment for reuse in individual sites.

Extending the core functionality of *Omeka S*, this grant proposes to more fully integrate digital collections in the LOD universe and to create new modes of access and dissemination through other platforms. This work will include software development, documentation, and support, and a set of developer training events.

A. Software Development

Resource Template Development: We will develop a **Basic Description Resource Template** to be shipped with the core *Omeka S* package that works to increase the chances that resource descriptions created within

[10] *Omeka* Dev list: https://groups.google.com/forum/#!forum/omeka-dev.

the platform will be easily sharable with key aggregators that make up the national digital platform by fulfilling the requirements of the DPLA MAP and the Europeana Data Model.

Module Development: We will scope, design, develop, and test the following four modules that enhance both description and dissemination of resources:

1. **Library of Congress Vocabularies Suggest:** adds an auto-complete feature to almost any metadata field in the *Omeka S* resource creation form by pulling results from LC's authorities and controlled vocabulary endpoints: http://id.loc.gov/;
2. **Getty Vocabularies Suggest:** adds an auto-complete feature to almost any metadata field in the *Omeka S* resource creation form by pulling results from the GRI Vocabularies' endpoints: https://www.getty.edu/research/tools/vocabularies/lod/index.html;
3. **Local Controlled Authorities:** allows an administrator to define a controlled vocabulary for any metadata field using a simple interface;
4. **Sharing and Social Media:** enables site visitors to share individual *Omeka S* resources as an iframe in other webpages, and through postings to popular social media platforms such as WordPress, Tumblr, Instagram, Facebook, Twitter, and Pinterest.

The *Omeka* team (*See Personnel below*) will build these templates and modules using the agile development process that we have implemented to good success with the initial *Omeka S* development. First, the development team meets to outline the requirements and specifications for code development. Then, a key developer (in the case of these modules, Jim Safley, with support from Patrick Murray-John and Kim Nguyen) takes the lead in building the initial iteration of the module to satisfy the specification. This initial development usually takes a number of two-week sprint cycles, depending on the complexity of the module. The full team meets at the end of each sprint to review the work accomplished during that period and to offer commentary and perhaps mid-course correction on the development strategy. The code developed for *Omeka S* integrates a set of automated unit tests that provide immediate feedback on the viability of that code as it is being developed.

Once, the development team has a functioning module that meets the requirements and specifications, they hand the module off to the quality assurance members of the team for testing. Megan Brett, Patrick Murray-John, and a graduate student will handle the human testing for this work. To do this testing, Brett and Murray-John will develop schema for internal and external testing of the modules. Brett and Murray-John will work with a development installation of the core software, which is populated with at least 50 resources, comprised of different media types with a range of descriptive data. They then make a fresh installation of the module and work through the key end-user functionality that was listed in the initial scope of work for the module. They test this functionality in environments that use the Mac OS, the Windows OS, and the most popular web browsers (Firefox, Chrome, Microsoft IE, and Safari). Once the testers are satisfied that the individual module functions as intended, they then test for compatibility with other modules. Along the way, the testers note in GitHub any possible bugs, formatting, or user experience issues. Based on the feedback reported in GitHub, the development team revisits the module code to correct any bugs or functional issues that arose in QA testing. This process of testing and correction continues through the sprint cycles until the entire team is satisfied with the product.

Then, the *Omeka* development team will package the module as a release candidate, which it will publicize through the *Omeka* Dev list. Members of the *Omeka* Dev list have a proven track record of offering invaluable feedback on *Omeka* extensions and we will gather that feedback before a final round of revision for each module. During the period that the members of the *Omeka S* development community are testing and exploring the modules, Brett and Murray-John will create the documentation and training materials necessary to support end-users and developers with a wide range of technical skills. Once the human testing has proven that the software is ready for release, the *Omeka* team will tag it in GitHub and package it for download from the *Omeka S* Modules directory. All code produced for *Omeka S* by RRCHNM is released with GPLv3 licensing.[11]

B. Developer training

Over the life of the *Omeka* family of products, the *Omeka* team has worked very hard to foster a strong developer community, through forums and discussion lists, providing detailed documentation, and publishing code on GitHub. *Omeka* competency continues to grow within the library community (a recent search for jobs on Code4Lib, shows 82 positions are tagged Omeka: http://jobs.code4lib.org/jobs/omeka/). Throughout the development phases of the *Omeka S* modules, members of the *Omeka* team will continue to cultivate relationships with the GLAM developer community and will respond to support forums and developers' lists.

Additionally, two team members will plan and conduct **four face-to-face workshops** that will help support developers in using and extending the development work done by the *Omeka* team. Training topics for the workshops will include an introduction to *Omeka S*'s implementation of JSON-LD, best practices for consuming data from *Omeka S*, module structure, and how to submit data back to *Omeka S*. More broadly, training will involve helping participants think about how *Omeka S* fits with the rest of the LOD cloud and techniques for leveraging those interactions. First, we will make proposals to conduct **workshops at the October 2017 Digital Library Federation (DLF) Forum, the February 2018 Code4Lib conference, and the June 2018 American Library Association (ALA) Annual Conference**. Then, RRCHNM will host a **workshop on the GMU Arlington campus in August 2018** that is open to developers in the DC area, particularly those working in federal libraries, archives, and museums, and any others interested in attending. Similar to the "playdates" we hosted early in *Omeka*'s original development, we will encourage a wide range of users and developers to come and learn about *Omeka S* and its modules. Each day-long workshop will target 25-30 developers and administrative end-users working in GLAMs to introduce them to the *Omeka S* platform and help prepare them to build, modify, or merely configure modules that will serve their digital project needs. In particular, we will make a concerted effort to reach out to under-represented communities of GLAM staff encouraging them to participate in the workshops (*See Diversity Plan below*). In sum, **we aim to train and support 100 developers** during the course of the grant.

4. Diversity Plan

In creating easy-to-use open source software, RRCHNM has always maintained a commitment to serving the needs of the widest community of users possible. As such, we make an effort to reach out to and support the work of under-represented and under-resourced populations. Our publicity strategy for the software beta testing and release, and for recruiting participants for workshops proposed here will continue that tradition. Specifically, we will take advantage of the vibrant Society of American Archivists (SAA) Roundtables (Archivists & Archives of Color Roundtable, Latin American and Caribbean Cultural Heritage Archives Roundtable, Lesbian & Gay Archives

[11] GPLv3: http://www.gnu.org/copyleft/gpl.html.

Roundtable (LAGAR), Native American Archives Roundtable, Public Library Archives/Special Collections Roundtable, and Women Archivists Roundtable); ALA Affiliate Organizations (American Indian Library Association, Asian/Pacific American Librarians Association, Black Caucus of ALA, Chinese American Librarians Association, REFORMA, and Urban Libraries Council); the ALA's Roundtables (Ethnic and Multicultural Information Exchange Roundtable, and Gay, Lesbian, BiSexual, and Transgender Roundtable), and the Association of Tribal Archives, Libraries, and Museums as points of contact and dialogue. We will publicize all software releases, workshops, and training materials with these constituencies, encouraging their members and colleagues to attend training events and make liberal use of the *Omeka* forums, the *Omeka* Dev list, and other resources.

5. Project Resources: Personnel, Time, Budget

Over the course of three years, the *Omeka* team will follow their proven agile development process to create the resource template, and the description and dissemination modules proposed above. The *Omeka* team has a history of creating robust, easy to use, open source software, as evidenced by the large user base for *Omeka Classic* and Omeka.net.

Personnel (*See Key Project Staff Resumes for more details on qualifications*):

- **Sharon M. Leon** (Primary Investigator) is *Omeka S*'s Project Director. She will be responsible for overall vision and direction of the project. Leon will dedicate 2% of her time to this work during the academic year and 10% of her time during the summer months.
- **Patrick Murray-John** is *Omeka S*'s Director of Developer Outreach. He will contribute to the software development, assisting Safley, and creating the Basic Description Resource Template and the Sharing and Social Media Module. Also, he will be responsible for creating developer training materials, and coordinating and co-teaching the developer workshops. Murray-John will oversee developer support during the grant. Murray-John will dedicate 15% of his time to this work.
- **James Safley** is a Software Developer for *Omeka S*. He will be responsible for the design and development of specifications for the software, for building the controlled vocabulary modules, and for responding to and resolving any issues that arise during the testing and release candidate stages of development. He will co-teach the developer workshops with Murray-John. Safley will dedicate 25% of his time to this work.
- **Kim Nguyen** is a Web Designer and *Omeka S*'s user experience and user interface designer. She will create all of the administrative user interfaces for the proposed modules and ensure that the public-facing functionality operates with the current *Omeka S* themes. Nguyen will dedicate 13% of her time to this work.
- **Megan Brett** is *Omeka S*'s End-User and Outreach Coordinator. She will be responsible for coordinating the quality assurance testing of the modules and creating the supporting documentation for the modules. Brett will also coordinate the publicity for the developer workshops. Brett will dedicate 15% of her time to this work.

We have included funding ($11,100) for travel for two members of the *Omeka* team (Murray-John and Safley) to three meetings (2017 DLF Forum; 2018 code4lib; 2018 ALA) to conduct workshops for developers. These funds will cover conference registration, airfare, hotel, and per diem. Additionally, we have included $500 for publicity materials, including stickers and bookmarks.

Work Plan: The *Omeka* team manages their work and communication through Slack with GitHub and other integrations. The work of this project will proceed in overlapping phases, under Leon's oversight (*See Schedule of Completion for more detail*).

- *Phase 1: Basic Resource Description Template* (October 1, 2015 to December 31, 2015 -- 3 months): Software design and development: Murray-John; Testing, documentation, and outreach: Brett.
- *Phase 2: Local Controlled Vocabulary Module* (October 1, 2016 to December 31, 2016 -- 15 months): Software design and development: Safley; User interface design: Nguyen; Testing, documentation, and outreach: Brett and Murray-John.
- *Phase 3: LC Vocabularies Suggest Module* (July 1, 2016 to June 30, 2017 -- 12 months): Software design and development: Safley and Murray-John; User interface design: Nguyen; Testing, documentation, and outreach: Brett and Murray-John
- *Phase 4: Getty Vocabularies Suggest* Module (July 1, 2017 to June 30, 2018 -- 12 months): Software design and development: Safley and Murray-John; User interface design: Nguyen; Testing, documentation, and outreach: Brett and Murray-John
- *Phase 5: Sharing and Social Media Module* (January 1, 2018 to June 30, 2018 -- 6 months): Software design and development: Murray-John; User interface design: Nguyen; Testing, documentation, and outreach: Brett and Murray-John
- *Phase 6: End-User and Developer Support* (January 1, 2016 to September 30, 2018 -- 33 months): Murray-John and Brett, with assistance from the rest of the *Omeka* development team
- *Phase 7: Developer Workshops* (October 1, 2017 to September 30, 2018 -- 12 months): Materials creation and evaluation: Murray-John; Outreach and publicity: Brett; Instruction: Safley and Murray-John (DLF, October 2017; code4lib, February 2018; ALA, June 2018; RRCHNM, August 2018)

6. Communications Plan

Communicating with *Omeka S* developers and end users is a crucial element of this grant's success. As such, these activities comprise a major portion of our deliverables. The *Omeka* team has a successful track record of reaching GLAM technologists and creating a welcoming environment for those who are new to the platform. The major venues for this communication will continue to be the *Omeka* forum and the *Omeka* Dev list. This allows us to address the needs of both end-users and developers. In overseeing this work, Murray-John will act as the Developer Outreach Lead, and Brett will serve as the End-User Outreach Lead. Leon as director of the project will address administrators and heads of libraries and archives as they consider the adoption and implementation of *Omeka S*. This work will be bolstered by the presence of strong end-user and developer documentation.

In addition to the open lines of communication available through the *Omeka* site, we will undertake a widespread publicity campaign. The *Omeka* blog and Twitter feed both reach thousands of followers, and we will use these outlets to offer progress reports and release updates about our work. At the same time we will make an effort to share our work with GLAMs and their staff who might be unfamiliar with *Omeka*, *Omeka S*, the description and dissemination modules under development, and the training workshops that we will offer in 2018. To reach these constituencies, we will provide targeted updates, news of releases, and invitations to our training workshops to dh+lib, the Association of College and Research Libraries Digital Humanities Discussion Group, the Public Library

Association, the SAA Student and New Archivist Roundtable, the SAA Metadata and Digital Object Roundtable, the Museum Computing Network, the College Art Association, and the American Alliance of Museums' Media and Technology Network. We will also provide similar updates to the venues listed in our Diversity Plan above.

7. Sustainability

Since 1994, RRCHNM has used digital media and computer technology to democratize history—to incorporate multiple voices, reach diverse audiences, and encourage popular participation in presenting and preserving the past. RRCHNM has established itself as a focal point for collaboration and communication within the emerging discipline of digital humanities with events such as the Rosenzweig Forum for Technology and Humanities and THATCamp, and the Press/Forward digital publishing initiative. The historians and staff at RRCHNM have created and collaborated on dozens of digital history projects, collectively making RRCHNM one of the most experienced digital humanities centers in the US. RRCHNM's work has been recognized with major awards and grants from the American Historical Association, the National Council on Public History, the National Humanities Center, the National Endowment for the Humanities, the Department of Education, the Library of Congress, the Institute of Museum and Library Services, the National Historic Publications and Records Commission, and the Mellon, Sloan, Hewlett, Getty, Kress, Rockefeller, and Kellogg foundations. RRCHNM has conducted more than $20 million in grant funded research, boasts an annual budget of nearly $2 million, and enjoys a nearly $3 million endowment (achieved with the assistance of two Challenge Grants from NEH), all of which guarantee its long-term stability.

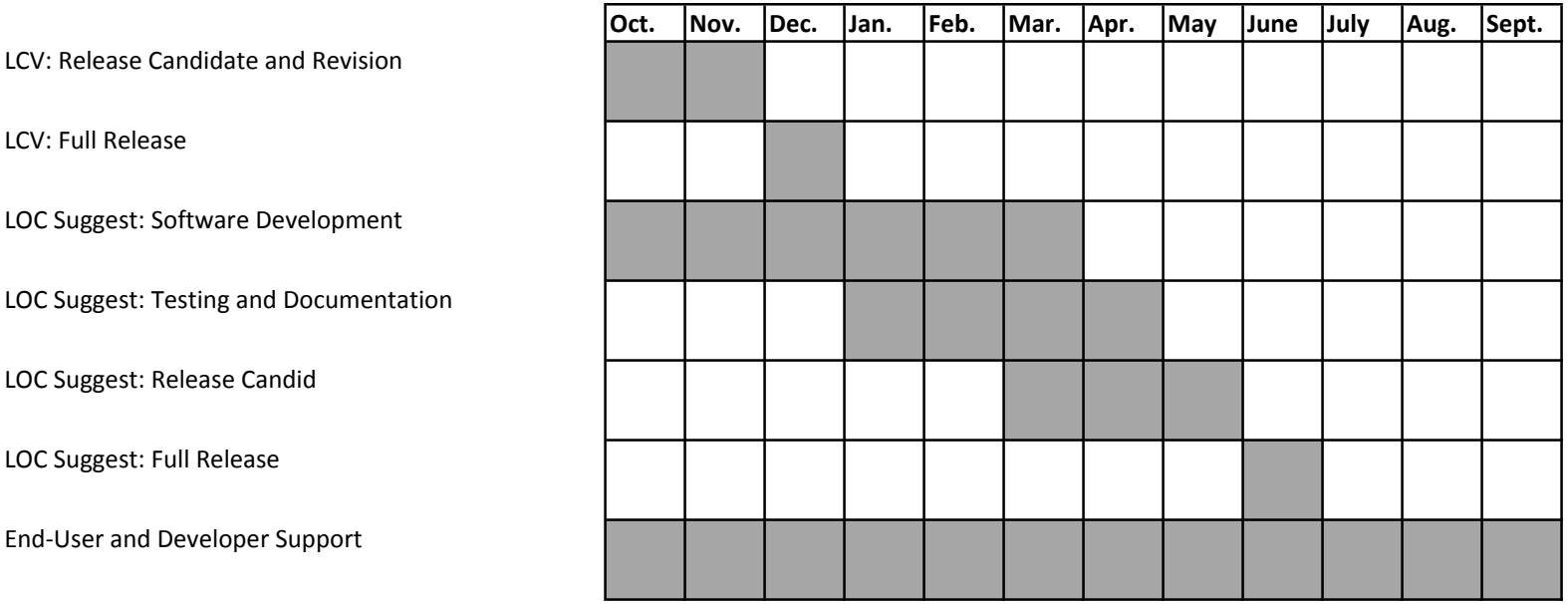
From the early stages of *Omeka* development, RRCHNM has fostered a strong open-source community that continues to thrive and to contribute back into the overall development and sustainability. In fact, *Omeka* was recognized by a blue-ribbon panel of leading technologists and Internet pioneers who awarded *Omeka* a Mellon Award for Technology Collaboration in December 2008, just ten months after the software's initial public launch. Today, over 630 developers participate in the *Omeka* Dev list and approximately 2,000 users post and answer less technical questions on the *Omeka* forums, which currently hosts over 18,000 posts. To encourage the community to dive in and fork the software, we moved *Omeka* to GitHub. Developers can follow the most recent code updates, comment on them, and submit bug reports. GitHub's social environment also encourages the community to share their unique branches of the core, plugins, or themes with others to use and test out themselves. *Omeka Classic* and *Omeka S* will continue to grow beyond RRCHNM.

Omeka.net is a key part of this sustainability strategy and its "freemium" fee structure will allow RRCHNM to cover the increased server space and maintenance costs required to run a hosted application. Financial and legal services and basic operations for Omeka.net are handled by the not-for-profit Corporation for Digital Scholarship (CDS) <http://digitalscholar.org/>, which returns net revenue from paid Omeka.net plans to RRCHNM for continued *Omeka* development and support. Currently, the revenue from Omeka.net supports one half-time developer and one half-time end-user support staffer for *Omeka* projects, and that support is tracking to increase in future years.

Schedule of Completion
Omeka S, Enhanced: Year 1

	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.
Basic Resource Description Template: Architecture Planning												
BRDT: Software Development												
BRDT: Documentation												
BRDT: Full Release												
Local Controlled Vocabulary Module: Architecture planning												
LCV: Software Development												
LCV: Testing and Documetation												
LOC Suggest Module: Architecture planning												
End-User and Developer Support												

Schedule of Completion
Omeka S, Enhanced: Year 2



Schedule of Completion
Omeka S, Enhanced: Year 3

	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.
Getty Suggest: Software Development												
Getty Suggest: Testing and Documentation												
Getty Suggest: Release Candidate and Revision												
Getty Suggest: Full Release												
Sharing and Social Media Module: Architecture and planning												
Social Media: Software Development												
Social Media: Release Candidate and Revision												
Social Media: Full Release												
End-User and Developer Support												
Workshops: Materials Development												
DLF Workshop												
code4lib Workshop												
ALA Workshop												
RRCHNM Workshop												

DIGITAL STEWARDSHIP SUPPLEMENTARY INFORMATION FORM

Introduction:
IMLS is committed to expanding public access to IMLS-funded research, data and other digital products: the assets you create with IMLS funding require careful stewardship to protect and enhance their value. They should be freely and readily available for use and re-use by libraries, archives, museums and the public. Applying these principles to the development of digital products is not straightforward; because technology is dynamic and because we do not want to inhibit innovation, IMLS does not want to prescribe set standards and best practices that would certainly become quickly outdated. Instead, IMLS defines the outcomes your projects should achieve in a series of questions; your answers are used by IMLS staff and by expert peer reviewers to evaluate your proposal; and they will play a critical role in determining whether your grant will be funded. Together, your answers will comprise the basis for a work plan for your project, as they will address all the major components of the development process.

Instructions:
If you propose to create any type of digital product as part of your proposal, you must complete this form. IMLS defines digital products very broadly. If you are developing anything through the use of information technology – e.g., digital collections, web resources, metadata, software, data– you should assume that you need to complete this form.

Please indicate which of the following digital products you will create or collect during your project.
Check all that apply:

	Every proposal creating a digital product should complete …	Part I
	If your project will create or collect …	**Then you should complete …**
☐	Digital content	Part II
☑	New software tools or applications	Part III
☐	A digital research dataset	Part IV

PART I.

A. Copyright and Intellectual Property Rights

We expect applicants to make federally funded work products widely available and usable through strategies such as publishing in open-access journals, depositing works in institutional or discipline-based repositories, and using non-restrictive licenses such as a Creative Commons license.

A.1 What will be the copyright or intellectual property status of the content you intend to create? Will you assign a Creative Commons license to the content? If so, which license will it be? http://us.creativecommons.org/

All code produced under the auspices of this grant project will be released as open source software under the GPLv3 license. All documentation and support materials will be released under the CC-BY 3.0 license.

A.2 What ownership rights will your organization assert over the new digital content, and what conditions will you impose on access and use? Explain any terms of access and conditions of use, why they are justifiable, and how you will notify potential users of the digital resources.

We will not be creating any new digital assets aside from the software, documentation, and support materials.

A.3 Will you create any content or products which may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities? If so, please describe the issues and how you plan to address them.

No.

Part II: Projects Creating Digital Content

A. Creating New Digital Content

A.1 Describe the digital content you will create and the quantities of each type and format you will use.

A.2 List the equipment and software that you will use to create the content or the name of the service provider who will perform the work.

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to create, along with the relevant information on the appropriate quality standards (e.g., resolution, sampling rate, pixel dimensions).

B. Digital Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

B.2 Describe your plan for preserving and maintaining digital assets during and after the grant period (e.g., storage systems, shared repositories, technical documentation, migration planning, commitment of organizational funding for these purposes). Please note: Storage and publication after the end of the grant period may be an allowable cost.

C. Metadata

C.1 Describe how you will produce metadata (e.g., technical, descriptive, administrative, preservation). Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

C.2 Explain your strategy for preserving and maintaining metadata created and/or collected during your project and after the grant period.

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content created during your project (e.g., an Advanced Programming Interface, contributions to the DPLA or other support to allow batch queries and retrieval of metadata).

D. Access and Use

D.1 Describe how you will make the digital content available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

D.2 Provide URL(s) for any examples of previous digital collections or content your organization has created.

Part III. Projects Creating New Software Tools or Applications

A. General Information

A.1 Describe the software tool or electronic system you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) the system or tool will serve.

Resource Template Development: We will develop a Basic Description Resource Template to be shipped with the core Omeka S package that works to increase the chances that resource descriptions created within the platform will be easily sharable with key aggregators that make up the national digital platform by fulfilling the requirements of the the DPLA Metadata Application Profile and the Europeana Data Model. Module Development: We will scope, design, develop, and test the following four modules that enhance both description and dissemination of resources: 1. Library of Congress Vocabularies Suggest: adds an auto-complete feature to almost any metadata field in the Omeka S resource creation form by pulling results from Library of Congress's authorities and controlled vocabulary endpoints: http://id.loc.gov/; 2. Getty Vocabularies Suggest: adds an auto-complete feature to almost any metadata field in the Omeka S resource creation form by pulling results from the Getty Vocabularies' endpoints: https://www.getty.edu/research/tools/vocabularies/lod/index.html; 3. Local Controlled Authorities: allows an administrator to define a controlled vocabulary for any metadata field using a simple interface; 4. Sharing and Social Media: enables site visitors to share individual Omeka S resources as an iframe in other webpages, and through postings to popular social media platforms such as WordPress, Tumblr, Instagram, Facebook, Twitter, and Pinterest. The intended audience for this work are medium to larger GLAMs wishing to create more standardize description for cultural heritage resources.

A.2 List other existing digital tools that wholly or partially perform the same functions, and explain how the tool or system you will create is different.

Currently, there are no modules for Omeka S that perform these functions. Several other curation/exhibit tools exist for cultural heritage work: Drupal, the Pawtucket module for Collective Access, and Spotlight for Hydra.

An LC Subject Headings Suggest, and a Social Media plugin exist for Omeka Classic and Omeka.net. This work will bring authority files, controlled vocabularies, and basic DPLA-Europeana compatible description to the Omeka S system, which is designed for medium to large libraries, and is built on a linked open data framework.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your new digital content.

Built using PHP 5.4 with Zend Framework 2, Doctrine, and Composer, Omeka S's Read/Write REST API enables all of the major software actions, such as the creation of users, items, sets, and sites. Omeka S uses JavaScript Object Notation-Linked Data (JSON-LD) as its native data format and ships with popular RDF vocabularies (DCMI Terms, DCMI Type, Bibo, and FOAF), which makes it possible to enmesh Omeka S in the LOD world.

B.2 Describe how the intended software or system will extend or interoperate with other existing software applications or systems.

Several elements make Omeka S and these modules extremely interoperable. First, the use of JSON-LD with RDF vocabularies primes Omeka S resources to for linking. Second, the existence of a REST API means that materials can easily be shifted in and out of the system. Finally, modules exist that facilitate the flow of materials from Fedora and DSpace repositories, and Zotero collections into the Omeka S environment for reuse in individual sites.

B.3 Describe any underlying additional software or system dependencies necessary to run the new software or system you will create.

The authority suggest modules will rely upon access to the Library of Congress and Getty Vocabulary endpoints cited above.

B.4 Describe the processes you will use for development documentation and for maintaining and updating technical documentation for users of the software or system.

During the process of development, the Omeka team provides developer documentation through Read-the-Docs. This is process that we have used successfully with the initial development of Omeka Classic: http://omeka.readthedocs.org/en/latest/. Additionally, there will be a set of documentation for end-users that follows the pattern of our work for Omeka Classic: http://omeka.org/codex/

B.5 Provide URL(s) for examples of any previous software tools or systems your organization has created.

Omeka Classic: http://omeka.org, https://github.com/omeka/
Omeka.net (Hosted): http://omeka.net
Scripto: http://scripto.org
Zotero: http://zotero.org

C. Access and Use

C.1 We expect applicants seeking federal funds for software or system development to develop and release these products as open source software. What ownership rights will your organization assert over the new software or system, and what conditions will you impose on the access and use of this product? Explain any terms of access and conditions of use, why these terms or conditions are justifiable, and how you will notify potential users of the software or system.

All Omeka software is released under the GPLv3.0 license.

C.2 Describe how you will make the software or system available to the public and/or its intended users.

We will distribute the products of this grant through the main Omeka site (http://omeka.org/). Additionally, all code will be freely available on GitHub (https://github.com/omeka/).

Part IV. Projects Creating Research Data

1. Summarize the intended purpose of the research, the type of data to be collected or generated, the method for collection or generation, the approximate dates or frequency when the data will be generated or collected, and the intended use of the data collected.

2. Does the proposed research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity already been approved? If not, what is your plan for securing approval?

3. Will you collect any personally identifiable information (PII) about individuals or proprietary information about organizations? If so, detail the specific steps you will take to protect such information while you prepare the research data files for public release (e.g. data anonymization, suppression of personally identifiable information, synthetic data).

4. If you will collect additional documentation such as consent agreements along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

5. What will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

6. What documentation will you capture or create along with the dataset(s)? What standards or schema will you use? Where will the documentation be stored, and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

7. What is the plan for archiving, managing, and disseminating data after the completion of research activity?

8. Identify where you will be publicly depositing dataset(s):

Name of repository: _____

URL: _____

9. When and how frequently will you review this data management plan? How will the implementation be monitored?

Appendix A
Omeka Project History

When the RRCHNM first applied for funding to support the development of *Omeka* (http://omeka.org) from Institute of Museum and Library Services (IMLS), few options existed for cultural heritage institution wishing to publish collections and narrative exhibits to the web easily. Targeted initially at small museums and historical societies, the system, which has been downloaded over 100,000 times since launch in 2008, has been widely adopted by a range of libraries, archives, museums, scholars, and popular users. *Omeka* provides a free and open source answer to the need for a web publishing platform that honored the importance of standards-based metadata and that allowed their content experts to showcase their unique knowledge about their collections, and allowed other experts, such as scholars, to better use these materials in their work. RRCHNM designed the software from its earliest stages to be interoperable with other content management systems, and item data is shareable through a variety of output formats, including RSS, Atom, JSON, and other XML feeds. With a few key plugins installed, users may make their data recognized by *Zotero*, a Mellon and IMLS-funded research tool, or import a *Zotero* library into an *Omeka* website. *Omeka's* most recent stable version is 2.3, which offers users a wide range of plugins to extend the basic functionality, a REST API, and a set of fully responsive themes so that *Omeka* sites render on mobile devices.

In October 2010, RRCHNM launched a public beta version of the hosted software, Omeka.net, to remove the barrier to adoption posed by the need for access to a Linux server environment and the need for staff comfortable working in that environment. This service offers any institution or individual the means to launch an *Omeka* site by removing the technological challenges of installation and maintenance, which is similar to cloud-based content management services offered by WordPress.com or Blogger. Omeka.net users—students, curators, faculty, librarians—are free to focus on building content from digital collections without fussing with server requirements, haggling with uncooperative IT departments, or waiting for budget allocations to pay for a web designer or developer. With Omeka.net, any user with an Internet connection can build digital exhibits, map photographs, collect memories from web audiences, or publish new scholarship in a few easy steps. The response to Omeka.net has been very promising and has become a popular platform for instructors in humanities disciplines who require undergraduate and graduate students to write and produce digital projects. Tutorials available in Omeka.net's Help section and in the popular digital humanities resource, the *Programming Historian* http://programminghistorian.org/lessons/, contributes to its adoption as a platform for scholarly communications. Omeka.net's user base grows steadily from year to year.

While a number of larger institutions have bought subscriptions for Omeka.net, the hosted version of the software does not offer IT staff at those institutions the degree of control that they would like. As a result, those technologists began to approach us asking for access to the codebase for Omeka.net, so that they could create and maintain similar centralized networks on their own servers. Comparing it to WordPressMU, libraries wanted to have one central administrative space to assign, upgrade, and manage *Omeka* websites for students, faculty, and staff to create digital portfolios, share research, and work collaboratively.

In October 2012, with support from the Andrew W. Mellon Foundation, the *Omeka* development team began work on a new version of the popular open source web-publishing platform to satisfy the needs of larger institutional users. *Omeka S* is the result of those efforts: a web publishing platform that allows a single administrator to offer a range of users—scholars, librarians, archivists, curators, students, teachers—the capacity to build sophisticated websites that employ cultural heritage collections owned by that institution, perhaps housed in a Fedora or DSpace repository, and in combination with materials available elsewhere on the web, such as items at Flickr and YouTube. This design offers IT staff the flexibility to

serve the needs of their content experts while not having to create massive data and bit stream redundancies by copying material from a repository system to the publishing platform, or storing local copies of cloud-hosted materials. More importantly, the design offers larger institutions a single point of administration for installation, software upgrades, and the extension of functionality and look and feel for all of the sites developed in the network. Together these features offer *Omeka S* administrators a critical balance of flexibility and control over their networks. (This flexibility and control is very important to *Omeka S* target audience of large institutional users, so they account for very few of our current Omeka.net subscriptions.)

To create *Omeka S*, the development team underwent a thoroughgoing process of rethinking the basic *Omeka* software architecture, guided by an initial code review conducted by Julie Meloni in late 2012. This required a willingness to leave behind some of the core infrastructural elements of the stand-alone version of *Omeka*, but the benefits of this open rethinking have been substantial. Built using PHP 5.3 with Zend Framework 2.0, Doctrine, Composer, *Omeka S*'s Read/Write Representational State Transfer (REST) Application Programming Interface (API) enables all of the major software actions, such as the creation of users, items, sets, and sites. *Omeka S* is also insulated from cross-site scripting and cross-site request forgery vulnerabilities, bolstering its security for deployment in an institutional setting.

Omeka S uses JavaScript Open Notation-Linked Data (JSON-LD) as its native data format, which makes it possible to enmesh *Omeka S* in the Linked Open Data world. Every *Omeka S* Resource (item, item set, media) has a URI, and the core software includes the following Resource Description Framework (RDF) vocabularies, which maximizes its data interoperability with other data publishers: Dublin Core Metadata Initiative (DCMI) Terms; DCMI Type; The Bibliographic Ontology (Bibo); and The Friend of A Friend Vocabulary (FOAF). In addition to including these popular LOD vocabularies, *Omeka S* offer users the ability to use the URIs for other *Omeka S* Resources as descriptive values within metadata fields, in essence linking one *Omeka S* Resource to another (i.e. using a Person type Resource for Thomas Jefferson as the value for the Creator field in description of the "Declaration of Independence" text Resource). *Omeka S* also makes it possible for users to attach media to Resources in three ways: through a simple file upload, through the use of an embed code from an outside Resource, or use of a URI for an existing Resource. Together, these features make prepare *Omeka S* to be fully embedded in the semantic web.

Finally, *Omeka S* integrates into the core software essential functionality that required plugins with the stand-alone version. Thus, all *Omeka S* instances will enable users to create sites with web pages that publish different types of content blocks, offering many different ways to view Resources. The result is flexible sites that are fully *Zotero* readable, and that capitalize on the improved core infrastructure that *Omeka S* offers.

Original Preliminary Proposal

ROY ROSENZWEIG Center FOR History AND New Media

Modules to Describe and Share *Omeka S* Resources

A Preliminary Proposal to the Institute of Museum and Library Services
January 20, 2015

Summary

The Roy Rosenzweig Center for History and New Media <http://chnm.gmu.edu> at George Mason University is requesting support for the *Omeka* team to create a set of five key modules for *Omeka S*, the new multisite version of the software that is targeted toward larger institutional users, that will facilitate the usage of a range of standardized and controlled vocabularies in the metadata creation process, and extend the sharing of those resources in other venues. This development will take approximately two years (October 1, 2015 to September 30, 2017) and will require approximately $220,000 in funds.

Background

Launched in February 2008, *Omeka* <http://omeka.org> has established itself as a leading open source web-publishing platform for digital collections. Targeted initially at small museums and historical societies, the system, which has been downloaded over 100,000 times since launch, has been widely adopted by a range of libraries, archives, museums, scholars, and popular users. *Omeka* provides a free and open source answer to the need for a web publishing platform that honored the importance of standards-based metadata and that allowed their content experts to showcase their unique knowledge about their collections, and allowed other experts, such as scholars, to better use these materials in their work.

In October 2012, with support from the Andrew W. Mellon Foundation, the *Omeka* development team began work on a new version of the popular open source web-publishing platform to satisfy the needs of larger institutional users. *Omeka S* is the result of those efforts: a web publishing platform that allows a single administrator to offer a range of users—scholars, librarians, archivists, curators, students, teachers—the capacity to build sophisticated websites that employ cultural heritage collections owned by that institution, perhaps housed in a DuraSpace repository, and in combination with materials available elsewhere on the web, such as items at Flickr and YouTube. This design offers IT staff the flexibility to serve the needs of their content experts while not having to create massive data and bit stream redundancies by copying material from a repository system to the publishing platform, or storing local copies of cloud-hosted materials. More importantly, the design offers larger institutions a single point of administration for installation, software upgrades, and the extension of functionality and look and feel for all of the sites developed in the network. Together these features offer *Omeka S* administrators a critical balance of flexibility and control over their networks.

To create *Omeka S*, the development team underwent a thoroughgoing process of rethinking the basic *Omeka* software architecture, guided by an initial code review conducted by Julie Meloni

in late 2012. This required a willingness to leave behind some of the core infrastructural elements of the stand-alone version of *Omeka*, but the benefits of this open rethinking have been substantial. Built using PHP 5.3 with Zend Framework 2.0, Doctrine, Composer, JQuery, and Symphony Console, *Omeka S's* Read/Write Representational State Transfer (REST) Application Programming Interface (API) enables all of the major software actions, such as the creation of users, items, sets, and sites. *Omeka S* is also insulated from cross-site scripting and cross-site request forgery vulnerabilities, bolstering its security for deployment in an institutional setting.

Omeka S uses JavaScript Open Notation-Linked Data (JSON-LD) as its native data format, which makes it possible to enmesh *Omeka S* in the Linked Open Data world. Every *Omeka S* Resource (item, item set, media) has a URI, and the core software includes the following Resource Description Framework (RDF) vocabularies, which maximizes its data interoperability with other data publishers: Dublin Core Metadata Initiative (DCMI) Terms; DCMI Type; The Bibliographic Ontology (Bibo); The Friend of A Friend Vocabulary (FOAF); and WGS84 Geo Positioning (Geo). In addition to including these popular LOD vocabularies, *Omeka S* offer users the ability to use the URIs for other *Omeka S* Resources as descriptive values within metadata fields, in essence linking one *Omeka S* Resource to another (i.e. using a Person type Resource for Thomas Jefferson as the value for the Creator field in description of the "Declaration of Independence" text Resource). *Omeka S* also makes it possible for users to attach media to Resources in three ways: through a simple file upload, through the use of an embed code from an outside Resource, or use of a URI for an existing Resource. Together, these features make prepare *Omeka S* to be fully embedded in the LOD universe.

The Need

The complete revision of the data model for *Omeka S* means that it is not possible to simply upgrade popular plugins from the *Omeka* stand-alone universe; the modules must be recreated from scratch. The scope of this task is not insignificant and will require significant developer resources, human testing, documentation, and support. Given that large institutions, particularly libraries, archives, and special collections are the primary user base for the *Omeka S* platform, RRCHNM proposes to build five key modules that will support data interoperability and enhanced distribution of collections:

1. **Library of Congress Subject Heading Suggest:** adds an auto-complete feature to almost any metadata field in your Omeka site by pulling results from Library of Congress's list of authorities and controlled vocabularies
2. **Getty Thesaurus of Geographic Names:** adds an auto-complete feature to almost any metadata field in your Omeka site by pulling results from the Getty's Thesaurus of Geographic Names
3. **Controlled Vocabulary:** allows an administrator to define a controlled vocabulary for any metadata field using a simple interface
4. **Embed Codes:** enables end-users to share individual Omeka items as an iframe in other webpages, blogs, and social media sites
5. **Custom Reports:** enables end-users to create custom edited collections that can be served as an HTML page or printed using a stylesheet