"RE:Search" - Unpacking the Algorithms That Shape Our UX
_____

Our technological experiences are increasingly mediated by algorithms - the code and computational processes embedded into our software. We, and our patrons, routinely engage in systems that predict, recommend, and speculate about our interests based on the digital fingerprint we provide with our link clicks and "likes," but we all struggle understanding how and why those systems work as they do. In the interest of building new competencies for our field, The Montana State University Library seeks $49,949.39 in funding from the Institute of Museum and Library Services through a Laura Bush 21st Century Librarian Program planning grant under the National Digital Platform category to support the *development of an Open Education Resource (OER) curriculum, pilot course, and teaching tool (a proof of concept search application)* to address the challenges of understanding and teaching about algorithms.

We are looking for ways to address a knowledge gap in our field: a lack of an understanding around the rules that govern our software and shape our digital experiences. Librarians have a long history of teaching digital literacy; however, while our ACRL Framework for Information Literacy considers how authority is constructed, our instruction programs have not kept up with how algorithms in software construct our experience. We are calling for a new competency that we have termed "Algorithmic Awareness". Librarians can apply this new competency to enrich our instruction programs and extend our relevancy by defining an emerging form of digital literacy for our patrons. At its core, this project is about introducing a new expertise into our profession and providing a new teaching moment for our librarians.

The planning grant will primarily support releasing this curriculum for teaching about algorithmic awareness; teaching a pilot course on algorithmic awareness for librarians; development of a working proof of concept application showing algorithms in practice; convening of an advisory council to shape the curriculum via online sessions and an in-person meeting; conducting testing and evaluation of librarian and library patron's awareness of algorithms and their role in our online experiences; and hiring of a research assistant to support grant work activities. To achieve the goals of this project, we will conduct work over 1 year in three main phases: *Testing, Evaluation, and Curriculum Development, Pilot Implementation,* and *Education.*

We see multiple opportunities to measure and understand the impact of our grant work including: the reuse or repurposing of the OER curriculum in other teaching settings; movement toward our OER course being picked up and taught inside of a library science or university curriculum; new public library classes or events based on our "Algorithmic Awareness" topic. etc. However, our overall goal is to find transparency for the invisible logic embedded in our software interactions. Success in this setting would be our community finding new teaching methods and confidence to make this logic visible for our patrons and ourselves.

"RE:Search" - Unpacking the Algorithms That Shape Our UX
_____

The Montana State University Library seeks $49,949.39 in funding from the Institute of Museum and Library Services through a Laura Bush 21st Century Librarian Program planning grant under the National Digital Platform category to conduct an environmental scan of the field's knowledge of algorithms, develop a proof of concept search application employing common algorithms, and create an Open Education Resource (OER) curriculum and pilot class that will be taught to librarians to improve our digital literacy around the algorithms that define our online experiences and shape our world.

---

1. Statement of National Need

1.1 Summary

Our technological experiences are increasingly mediated by algorithms - the code and computational processes embedded into our software. Recent work by scholars, such as Dr. Safiya Umoja Noble, has shown [how algorithms exhibit implicit biases and reify societal prejudices](). Moreover, the technical nature of algorithms and the lack of transparency surrounding them can be a challenge for novices. We, and our patrons, routinely engage in systems that predict, recommend, and speculate about our interests based on the digital fingerprint we provide with our link clicks and "likes", but we all struggle understanding how and why those systems work as they do. In seeking to understand common systems, like the Facebook news feed or the Google search engine results page, we view this grant research as an opportunity to discover the scope and reach of algorithms and how they might be taught. We are looking for ways to address a gap in our field: a lack of an understanding around the rules that govern our software and shape our digital experiences. Librarians have a long history of teaching digital literacy; however, while our [ACRL Framework for Information Literacy]() considers how authority is constructed, our instruction programs have not kept up with how algorithms in software construct our experience. We are calling for a new competency that we have termed "Algorithmic Awareness". Librarians can apply this new competency to enrich our instruction programs and extend our relevancy by defining an emerging form of digital literacy for our patrons. At its core, this project is about introducing a new expertise into our profession and providing a new teaching moment for our librarians.

1.2 Algorithms, Education, and Libraries: A Brief Introduction

For our purposes, we are defining algorithms in the terms of how they are applied within computer programs. In this context, an algorithm is an "instance of logic written in software by software developers to be effective for the intended "target" computer(s) to produce output from given (perhaps null) input" ([Wikipedia article on Computer Algorithms](), 2017). Our focus is in how these logical rules are explicated and can be brought into a the library pedagogical environment.

In our literature review, we noted increasing calls to consider the algorithm as an object of study. What was striking in this review was that most of the earlier studies look at the efficacy or construction of algorithms. This focus on the practice of building algorithms is seen in many studies featuring the benchmarking or comparison of algorithms in practice (dos Santos et. al 2016, Kibekbaev et. al. 2017). However more recently, the research takes a turn towards the social impacts of algorithms. Safiya Umoja Noble was leading this discussion when she asked: "What kinds of results do Google's search engine provide about Black girls when keyword searching?" (Noble 2013). These questions referencing the deleterious effects of common algorithms are picked up again by Michelle Willson when she considers the place of algorithms in our everyday online experiences (Willson 2017). Even further, these questions are given a vocabulary when Taina Bucher introduces her concept of "the algorithmic imaginary - ways of thinking about what algorithms are, what they should be and how they function" in looking at people's experience of the algorithms present in Facebook (Bucher 2017). Bucher's concept of the "algorithmic imaginary" maps directly into our ideas around an "Algorithmic Awareness" competency for librarians. And finally, we were particularly interested in Rob Kitchin's study where he notes the "three main challenges that hinder research about algorithms (gaining access to their formulation; they are heterogeneous and embedded in wider systems; their work unfolds contextually and contingently), which require practical and epistemological attention." We agree with Kitchin's formulation and see our project as an answer to his call for "practical and epistemological attention". We also see the design of our project with our emphasis on algorithms in a prototype search application where we can "gain access to their formulation" as a means to bring some transparency to how the algorithms work in practice.

In the interest of understanding the learning landscape, we also conducted a survey of top Library and Information Science curriculums which showed that information literacy as a concept is offered. Increasingly, more and more technical topics - data science and analytics, information management and security, web design and computer programming - appear in these course catalogs, but the treatment and discussion of algorithms as objects for study or as a part of information literacy is not present at the time of this writing. All of the Library Science curriculums offer an "Advanced" or "Special Topics" course which would accommodate the complete curriculum we are looking to create. In our look at academic libraries, we noted instruction topics based on academic tools like citation management software or beginning classes on information literacy. Digital literacy is a rising topic and would suggest our abridged curriculum will have a home in these academic library settings as well. Similarly, a survey of public library instruction topics shows a focus on traditional literacy and computer skills/training and increasingly an interest in privacy in the digital environment, but the discussion of "Algorithmic Awareness" is not yet present. In spite of this, our takeaway is that many public libraries offer instructional efforts or classes that will provide a forum for the abridged curriculum we are proposing. An abridged sample of the the resources we surveyed and a complete bibliography appears in Section N, Supplemental Materials.

1.3 Project Overview: An OER Curriculum, Teaching Tool, and Course for Teaching Algorithms

We propose the *development of a curriculum, pilot course, and teaching tool (a proof of concept search application)* to address the challenges of understanding algorithms. While many librarians agree that librarians should understand the workings of the software we use as a form of digital literacy, the application of algorithms in our digital experiences and the lack of educational resources around these rules that govern our software interactions demonstrate that the library profession needs more resources, more dialogue, and more direction for addressing the issue of algorithm design and implementation. This planning grant and the primary grant deliverables of a curriculum, a pilot course, and a proof of concept search application will prove vital in advancing our professional discourse, knowledge, and practice regarding algorithms and their place in our information literacy frameworks and train our librarians in a new competency we are calling "Algorithmic Awareness".

1.4 Relevance to Project Category and IMLS Goals

In developing a curriculum, a proof of concept, and a pilot course, the grant work promises to enhance our professional understanding of algorithms, increase our knowledge and capacity with digital library tools, and expand the profession's information literacy competencies. Additionally, this outcome aligns with the IMLS agency theme of continuous learning in response to advances in technology. Key grant outputs—an OER curriculum, a pilot course, teaching tools, and open access publications—align with the IMLS agency theme of creating documentation and professional development that enables others to implement or replicate project outcomes. Furthermore, this project aligns with and will be measured using the IMLS' agency-level goal of learning and specifically the performance goals to "Train and develop library and museum professionals", "Support communities of practice", and "Develop and provide inclusive and accessible learning opportunities."

2. Project Design

2.1 Structure and Phasing

To achieve the goals of this project, work will unfold in three three main phases: *Testing, Evaluation, and Curriculum Development, Pilot Implementation,* and *Education.* Each phase will take 4 months to complete with work overlapping as outlined in the project work schedule. A detailed description of the activities in each phase are as follows:

2.1.1 Phase 1: Testing, Evaluation, and Curriculum Development

The core of the grant work involves developing an OER curriculum for teaching and learning about algorithms. An initial activity here will be to survey the field to gain a sense of the current knowledge about algorithms and interest in learning more about algorithms. This will give our grant work a baseline and provide us with a context for where to begin our educational efforts. The curriculum itself will be developed in a modular format with an initial, extended format around a semester-length course. Two abridged formats for the course will also be developed in tandem to allow for a multi-day instruction (week-long) and a single-day course. The goal is to allow for multiple teaching and training experiences featuring the highlights and essential

takeaways derived from the extended course format. Topical decisions and teaching methods will be discussed and determined by the advisory council during this phase. Toward the end of this phase, the council will draft and finalize the OER curriculum with the goal of placing the class into a public teaching venue or workshop setting. The OER curriculum will be released on our MSU Library GitHub account alongside the code release for our teaching tool mentioned below in section 2.1.2.

More specific tasks in this phase are itemized here and include: Drafting a curriculum on "Algorithmic Awareness" with an emphasis on training instructors to carry forward what they have learned into future instruction. Bringing together an advisory group including scholars, librarians, administrators, and designers to shape the curriculum. Preparing and conducting an environmental scan of algorithm awareness in the profession using interviews and surveys. Usability testing of selected library patrons with an emphasis on how current systems like Twitter, Facebook, and Google use algorithms as part of their user experience. Releasing of a white paper explaining how algorithms work including common, generic algorithms in use by popular systems such as Twitter, Facebook, and Google.

2.1.2 Phase 2: Pilot Implementation for Teaching Tool (Proof of Concept Search Application)

As a frame of reference for algorithms, this project will look at how search algorithms work in practice. Search remains a key component of the library experience. What is often misunderstood are the digital machinations, and the ghosts in the [search] machine that produce search results for our users. To this end, MSU Library has an existing search prototype and code release on GitHub that can provide the foundation for this phase of work. The additional work scoped for the grant will include refinements for annotation and a new release of this "search-ux" tool as an open-source teaching tool in our MSU Library GitHub account. During the grant period, we will also be testing the user experience of the search tool, the utility of the annotations, and the best way to bring transparency about algorithms into the search experience. The goal here will be allow librarians to demonstrate common search algorithms, but give them and their learners a chance to slow down the process and highlight what is happening when a search facet is applied or a natural language query is interpreted by the machine. Some of the algorithms we will study and demonstrate include: PageRank, merge sort and heap sort, Dijkstra's algorithm, link analysis, and TF-IDF (Term Frequency-Inverse Document Frequency). We will use the search application to demonstrate and unpack the various algorithms present in the proof of concept as part of the teaching sessions. In the process, students will gain a perspective on how code becomes an interaction and we can establish how theory leads into practice.

More specific tasks in this phase are itemized here and include: Development of a proof of concept search application that applies common algorithms for relevancy, recommendations, understanding sentiment, linking relationships between users, and grouping documents into

categories. Proof of concept code released as an open source project for additional learning opportunities.

2.1.3 Phase 3: Education - Conduct Pilot Class, Post-Forum Activities, and Project Communication

Our supporting partners (ACRL, LITA, and DLF) featured in our letters of support have invited us to conduct the pilot course within their teaching venues. This phase will mark the implementation and testing of our curriculum. Project director, Jason Clark, will teach the first version of the course in the summer of 2018 within one of the partner's teaching venues. We will also see the final release of our OER curriculum during this phase. The advisory council will debrief and finalize project documentation. Key project staff will travel to the DLF 2018 Forum to present on the project and build awareness around the OER curriculum and the lessons learned during the grant work.

More specific tasks in this phase are itemized here and include: Conduct class using the "Algorithm Awareness" curriculum in a national forum including: the Digital Library Federation (DLF) National Forum or online as part of the Library Information Technology Association (LITA) Continuing Education program. Assess course learning objectives based on response of pilot class. Strategize about curriculum expansion by bringing curriculum to the Montana State University Library Curriculum committee for a potential semester-long course and seeking potential opportunities for the course within Library and Information Science graduate programs. Final (iterative) release of "Algorithmic Awareness" curriculum syllabus as an OER based on findings from teaching the pilot course.

2.2 Project Timeline

Phase	Duration	Activities
1. **Testing, Evaluation, and Curriculum Development**	November 2017-April 2018	-Advisory Board Meeting (March 2018)
2. **Pilot Implementation for Teaching Tool**	January 2018-May 2018	- User Experience Testing (February 2018)
3. **Conduct Pilot Class, Post-Forum Activities, and Project Communication**	May 2018-October 2018	- DLF Fall 2018 - ACRL Online Learning - LITA Online Course - ALA Summer 2018

2.3 Assumptions and Risks

In proposing this grant work, our primary assumption is that questions around algorithms and the ethics and rules that govern our software will continue to resonate within the library community throughout the duration of the grant. We furthermore assume that calls for transparency in how our digital experiences are mediated by decisions we make while building our software will continue to resonate and grow. For this reason, we assume that the outputs from this grant work will be of lasting value to the library community. We do have contingency plans for staffing changes within the Advisory Council as we have actively promoted the idea during the drafting of the grant and have received several positive responses from others in the profession who are interested in a role in the grant work. We also have testing planned for human subjects related to the user experience of the search tool prototype. There are no anticipated risks to participating subjects, and the following mechanisms will be put in place to ensure participant privacy: user experience data will be linked to participants through a secret key available only to the researchers via a secure system, research data will be stored securely and in confidence, and the final dataset will be de-identified.

3. Statement of National Impact
3.1 Outputs
This grant work will produce key deliverables that will offer immediate professional development advances for knowledge of algorithms and how to teach this gained competency. These deliverables include: an OER curriculum for teaching algorithm awareness; an open source proof of concept search application demonstrating how algorithms move from theory into practice; and a pilot course for librarians to learn how to recognize, teach, and decode algorithms for information literacy instruction. This project will be measured using the IMLS' agency-level goal of learning and specifically the performance goals to "Train and develop library and museum professionals", "Support communities of practice", and "Develop and provide inclusive and accessible learning opportunities."

The realization of these goals will ultimately produce several key outputs that will offer immediate practical and scholarly advances for understanding and teaching about algorithms:

- An **OER curriculum on Teaching Algorithms** that provides an educational blueprint for the pilot course and additional teaching opportunities with final synthesis and production by the Advisory Board, Jason Clark, and the Research Assistant (see Section 5.1, Personnel, for more).
- A **search tool prototype** that provides annotations and explanations as people are interacting with the search explaining that provides background, resources, and best practices to guide libraries in responsibly implementing web analytics. The code will be released as open source on the MSU Library GitHub account and licensed under the MIT License to allow for wide reuse and dissemination. The search tool prototype will be informed by Advisory Board input, with final synthesis and production by Jason Clark and the Research Assistant (see Section 5.1, Personnel, for more).

- A **pilot course on Algorithms** delivered by Jason Clark informed by advisory council input with final synthesis and production by Jason Clark (see Section 5.1, Personnel, for more).
- **A white paper, conference presentation** and/or **peer-reviewed publication** that will further disseminate planning activities and generate engagement around the teaching and understanding of algorithms. These outputs will be distributed openly on organizational websites and blogs (e.g., ACRLog, the DLF News, and the LITA website) and submitted to journals that are fully Open Access (sometimes called Diamond or Platinum OA) so as to allow for the widest possible dissemination at the lowest cost.

The grant work outputs will offer actionable products that can impact policy and practice so that the library community can translate its new knowledge around algorithms into the work of digital and information literacy .

3.2 Areas of Impact

Our project establishes a new competency of "Algorithmic Awareness" and grounds a new question for the information literacy initiatives at the center of our profession's instructional goals. In short, our profession will benefit from an intensive, focused study of algorithms and their role in mediating our online interactions that will ultimately lead to an actionable teaching agenda for learning and teaching about these invisible pieces of logic inside the machine. This work also brings awareness to a topic that is emerging as a new competency for citizens who live inside the filter bubble of our online systems. Additionally, The grant work aims to impact a wide spectrum of libraries: public, academic, community college, tribal, school, and special. Our first goal is to empower librarians in all settings with a new information literacy competency. In order to reach across this spectrum, the dialogue and outputs from the grant will be targeted at two main areas of impact: practice and policy. In terms of practice, the grant will produce a curriculum and course that will inform and shape practice by helping librarians teach algorithms to themselves and their communities. In terms of policy, we have built depth and diverse representation in our advisory group to include technologists, public librarians, academic librarians, and organizational heads. In this formation, we have a set of experts that can offer broad perspectives, resources, and access to venues (e.g., online teaching spaces) for turning our ideas into practice. We have also brought in library adjacent partners like Mozilla, the proudly non-profit champions of a healthy Internet, to extend the reach of our ideas and connect with an international technology community. Additionally, our support letters from our major professional organizations like ACRL, LITA, and the DLF demonstrate a professional commitment to our grant idea. With ACRL in particular, we see the potential for our curriculum, tools, and teaching practices to be enacted in the ACRL Framework for Information Literacy for Higher Education. All of these partnerships point to strong collaborations that can have an impact on the practice and policies around digital literacy in the library profession.

3.4 Measuring Impact

We see several means to understand and quantify the reach and impact of our project. Our first measure will be a successful running of the pilot course. Genuine interest and attendance in the

course will show that our idea has merit and has had potential impact in developing new teaching competencies for librarians. We will also survey students to ask for ways to improve the course and how they plan to use and apply what they have learned. A second measure will be the release of our proof of concept search application using an open license on our institutional GitHub account. We will be able to monitor engagement analytics using common metrics from the GitHub platform like stars, pull requests, or issues. A third measure will be a the release of an OER curriculum within our institutional GitHub account. We can use the same engagement analytics here as mentioned above. A fourth measure will be the reuse or repurposing of the OER curriculum in other teaching settings. Kate Eppler, an advisory council member has a strong interest in bringing the abridged curriculum to the learning opportunities at San Francisco Public Library. This repurposing of curriculum content will show growth in the idea and potential for reuse in other teaching settings. A fifth measure will be the acceptance of the idea into a presentation or publication forum. We have built this activity into the grant and our participation in a scholarly channel will confirm that the grant work is being discovered and valued by the librarian community. A final measure we will use is to see movement toward our OER course being picked up and taught inside of a library science or university curriculum. Project Director, Jason Clark, has several leads for potentially teaching this course in a university setting and placement inside of a library school would confirm that the grant work has had an impact.

3.4 Dissemination

The core activity in our dissemination plan is to conduct the course in a public venue. In addition, key project staff will review and synthesize our findings during curriculum and prototype development, ultimately producing: a blog post announcing the findings and the course on the DLF News site or ACRL Blog, a conference presentation, and a peer-reviewed publication. More specifically, key project staff will travel to relevant library and technology conferences for the purposes of generating dialogue around "Algorithmic Awareness" in libraries and communicating the activities and progress of the grant work. At each of these conferences, the key project staff will engage with the library community through presentations, workshops, and/or working groups that meet and discuss topics relevant to grant work. We have strategically planned travel so that the key project staff will be able to engage with diverse stakeholder communities at the DLF Fall 2018. Final project outputs will be delivered at DLF Fall 2018, which will coincide with the end of the grant period.

4. Diversity Plan

Our project grant did not require a formal diversity plan, but we have taken steps to keep the work and grant outputs inclusive and diverse. First, our advisory board participation was formed with broad representation from different sectors of librarianship and with an eye toward a depth of diversity. Second, our OER curriculum will allow all members of the library community to participate in this learning opportunity and reuse the course materials in any educational setting. Third, in setting expectations for our class and contributions to our software, we will include a code of conduct that guides behavior and offers a means of

reporting on bad actors within the communities associated around our teaching and coding environments. Helping to guide our work at inclusion within our curriculum and its associated activities, we have access to and will consult with members of Montana State University's University's Cultural Attunement Committee, which was created by [ADVANCE](Project) [Project](TRACS), a multi-year grant project funded by the National Science Foundation that aims to increase student and faculty diversity through research, policy, and training.

5. Project Resources
5.1 Personnel

Jason Clark, Associate Professor and Head of Library Informatics and Computing at Montana State University, will serve as Project Director and lead worker on grant activities. Jason builds and directs the digital branch of the Montana State University (MSU) Library. His work focuses on Semantic Web development, digital library development, metadata and data modeling, web services and APIs, search engine optimization, and interface design.

Research Assistant, This position will be filled by a Montana State University student and will work many of the primary activities on the grant including: collecting and synthesizing interview and user testing data; preparing project documentation; drafting written outputs and presentations related to dissemination efforts; and assisting with development/design of the curriculum and search application proof of concept. *(Hiring contingent on grant funding.)*

Scott Young, Assistant Professor and Digital Initiatives Librarian at Montana State University, will serve as a lead consultant for testing and user experience assessment. As a UX researcher and practitioner, Scott is highly attuned to the needs and expectations of users. Scott publishes, presents, and teaches on user experience, privacy, and ethics.

Lisa Janicke Hinchliffe, Professor and Coordinator for Information Literacy Services and Instruction in the University Library at the University of Illinois at Urbana-Champaign, will serve as lead consultant for instructional models. Lisa has presented and published widely on information literacy, teaching and learning, the value of academic libraries and library assessment, evaluation, innovation, analytics, and privacy.

5.2 Advisory Council
We have formed an advisory council to help guide the grant efforts. Each of the council members is an accomplished member of the library and technology professions and brings a wealth of knowledge to our project. The advisory council will meet quarterly to draft documentation, build the curriculum, offer feedback on the pilot software, and draft grant outputs as necessary. (This information is also itemized in the Key Staff Project Staff document under Section I.)

- [Lisa Janicke Hinchliffe](), Professor/Coordinator for Information Literacy Services - University of Illinois; member of ACRL's Information Literacy Immersion Program
- [Scott Young](), Digital Initiatives Librarian - Montana State University

- [Abigail Cabunoc Mayes](#), Lead Developer - Open Source Engagement at Mozilla
- [Kate Eppler](#), Program Manager & Learning Coordinator, The Bridge at Main - San Francisco Public Library
- [Jan Zauha](#), Professor, Reference and Outreach Librarian - Montana State University
- [Bethany Nowviskie](#), Director of Digital Library Federation (DLF)
- [Andromeda Yelton,](#) President of Library and Information Technology Association (LITA)

5.3 Budget

Our request for $49,949.39 will primarily support curriculum development and teaching, a research assistant, development of the search proof of concept, testing/evaluation, and travel. Salary requests would be used towards: a buy-out of time for the PI's work in coordinating grant work, teaching the curriculum in a pilot course, and developing the proof of concept, and a research assistant's work in analyzing interview data, preparing project documentation, and assisting with development/design of the curriculum and proof of concept. Direct costs include: $9,448.40 to cover 10% of the PI's time for 12 months based on annual salary; $14,180.40 to hire a research assistant at a rate of $13/hr for 1080/hrs; $2,675.00 for testing and evaluation services including incentives for participants, software, and tools for conducting the study; and $3,800 to cover travel costs for Project Director Clark's and his research assistant's dissemination of the project's outcomes at the 2018 Digital Library Federation Forum. Additional costs of $9,500 will allow the project advisory group to meet and attend Montana State University in March of 2018 to share project progress, generate dialogue, and create project outputs such as - curriculum, white paper, input on the search application, etc. - for curriculum development and algorithmic awareness in libraries. Indirect costs would be at 34.5% of modified total direct costs totaling $10,385.81 For full description and justification of the budget, please see Section G, Budget, and Section H, Budget Justification.

6. References

For a complete list of references, please see the bibliography included in Section N, Supplemental Materials.

Year 1	2017		2018									
	November	December	January	February	March	April	May	June	July	August	September	October
Phase 1: **Testing, Evaluation, and Curriculum Development**					Board of Advisors Meeting							
Phase 2: Pilot Implementation				User Experience Testing								
Phase 3: Education								ACRL Online Learning Course and LITA Online Course				DLF Forum

DIGITAL PRODUCT FORM

Introduction
The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (i.e., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. However, applying these principles to the development and management of digital products can be challenging. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

Instructions
You must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

PART I: Intellectual Property Rights and Permissions

A.1 What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

MSU Library makes no claims of ownership on the intellectual property of the content. The OER curriculum and course materials will be released under the Creative Commons CC-BY license. In addition, software will be released under the open source Apache License, Version 2.0 at http://www.apache.org/licenses/LICENSE-2.0.html.

A.2 What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

MSU Library makes no claims of ownership on the intellectual property of the content.

A.3 If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

We have testing planned for human subjects related to the user experience of the search tool prototype. There are no anticipated risks to participating subjects, and the following mechanisms will be put in place to ensure participant privacy: user experience data will be linked to participants through a secret key available only to the researchers via a secure system, research data will be stored securely and in confidence, and the final dataset will be de-identified.

Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

A. Creating or Collecting New Digital Content, Resources, or Assets

A.1 Describe the digital content, resources, or assets you will create or collect, the quantities of each type, and format you will use.

We will be creating textual content. The first content will be two documents: an OER curriculum and course materials in word processing document files (.docx) and in Markdown (.md) files. The second content will be two spreadsheet (.csv) files that willbe generated during user testing experiments.

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

Microsoft Office Suite (Word, Excel, etc.) for the office documents. SublimeText or Vim to create the Markdown (.md) files.

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

Microsoft Office Suite (Word, Excel, etc.) for the office documents which will be .docx or .xslx files. SublimeText or Vim to create the Markdown (.md) files.

B. Workflow and Asset Maintenance/Preservation

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

Workflow and storage integrity will be monitored weekly using our local systems staff and their quality control processes. Advisory council members and the research assistant will monitor and create a document organization structure using local storage options. Versioning of documents will be handled using standard office conventions present in MS Office software.

B.2 Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

All MSU Library projects are added to our web archiving and preservation servers which include multiple RAID levels (each partition capable of storing over 1TB of data) for distributed storage and allowing for multiple distributed copies. Once a project is brought into our development server the project is tracked and versioned in our local implementation of the Git software. All digital assets produced during the grant project will become part of this storage and asset management infrastructure, migrated to newer storage as needed, and absorbed into costs that are already part of the library operations budget for as long as the organization exists.

C. Metadata

C.1 Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata. 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).

We don't have plans to deposit this content in our institutional repository, ScholarWorks. But, if we did make that decision we would use the Dublin Core metadata standard. We will keep this in mind as we are working through the grant.

C.2 Explain your strategy for preserving and maintaining metadata created or collected during and after the award period of performance.

Not applicable

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

Not applicable

D. Access and Use

D.1 Describe how you will make the digital content, resources, or assets 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).

OER curriculum and course materials will be distributed and openly available on our institutional GitHub account - https://github.com/msulibrary.

D.2 Provide the name(s) and URL(s) (Uniform Resource Locator) for any examples of previous digital content, resources, or assets your organization has created.

Digital Library Content (full list at http://www.lib.montana.edu/digital/):
- Acoustic Atlas - http://acousticatlas.org/
- Range Science Information System - http://arc.lib.montana.edu/range-science/
- BookMeUp (Book recommendation tool) - http://www.lib.montana.edu/labs/bookme/
- MSU Historical Photos Walking Tour App - http://arc.lib.montana.edu/msu-photos/app/

Part III. Projects Developing Software

A. General Information

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

MSU Library has an existing search prototype and code release on GitHub that will provide the foundation for a new release of this "search-ux" app as an open-source teaching tool. The functions here will be allow librarians to demonstrate common search algorithms, but give them and their learners a chance to slow down the process and highlight what is happening when a search facet is applied or a natural language query is interpreted by the machine.

A.2 List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

In our research,we are not aware of another software that performs these functions. The strength of our implementation will be how the software makes the logic behind search actions transparent by offering annotations with the UI that explain what is happening.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

Our software development group at MSU Library works within the LAMP (Linux, Apache, MySQL, and PHP) stack. We will apply these technologies along with HTML5, CSS3, and Javascript to create our "search tool" teaching app. We chose these technologies due to our local expertise and we are looking for our tool to be a Web deliverable and HTML and Javascript are the cornerstones of the Web platform.

B.2 Describe how the software you intend to create will extend or interoperate with relevant existing software.

As we mentioned, the MSU Library has a search application (https://github.com/jasonclark/search-ux) that we are refining and extending for the teaching environment. The focus will be on new features that show the algorithms "in action" as people move through the search UI.

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

An Apache Web Server with PHP installed are the core dependencies for the search application.

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

All development and documentation will be version-controlled using Git. Code and supporting documentation will be packaged together.

B.5 Provide the name(s) and URL(s) for examples of any previous software your organization has created.

MSU Library Software (https://github.com/msulibrary)
- https://github.com/msulibrary/linked-people
- https://github.com/msulibrary/bib-template-fiction
- https://github.com/msulibrary/msu-research-citations

C. Access and Use

C.1 We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

Software will be released under the open source Apache License, Version 2.0 which is a permissive license similar to the MIT License, but also provides an express grant of patent rights from contributors to users. We have no problems in releasing a beta version of the software for reuse and application at the end of this grant work.
http://www.apache.org/licenses/LICENSE-2.0.html

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

Release on our MSU Library GitHub account - https://github.com/msulibrary

C.3 Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository: Montana State University Library on GitHub

URL https://github.com/msulibrary

Part IV: Projects Creating Datasets

A.1 Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

We intend to collect survey data and user testing data. Survey data will be collected in the first phase (November 2017) of the grant to get a baseline understanding of the library professions' knowledge of algorithms. User testing data using an interface walkthrough and series of tasks will be collected during the second pahse of the project (February 2017).

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

Our lead user experience consultant (Scott Young) on the grant and the project director (Jason Clark) are certified for user testing by the university IRB committee. A proposal for additional research activity will be put forward if additional testing becomes necessary.

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

There are no plans for collecting PII from participating subjects, but the following mechanisms will be put in place to ensure participant privacy: user experience data will be linked to participants through a secret key available only to the researchers via a secure system, research data will be stored securely and in confidence, and the final dataset will be de-identified.

A.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.

Consent forms will combined as a supplemental file to the data and linked using an administrative metadata record.

A.5 What methods 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).

These datasets will be simple, small .CSV files. Technical requirements will be minimal and can follow standard storage routines for textual documents.

A.6 What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? 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?

Not applicable.

A.7 What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

Our plan is to store the data in our local institutional repository, ScholarWorks. It is an archival platform and a public resource indexed in commercial search engines for good discoverability.

A.8 Identify where you will deposit the dataset(s):

Name of repository: MSU Library ScholarWorks

URL: http://scholarworks.montana.edu/

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

MSU Library has a well-established "Data Services" program (https://www.lib.montana.edu/services/data/) and a Data Management Librarian who looks at data management planning and the status of data deposits on a six month cycle. Data generated during this grant project would fall under these guidelines and maintenance schedules.