Library Values & Privacy in our National Digital Strategies: Field guides, Convenings, and Conversations

Abstract

The UW-Milwaukee Center for Information Policy Research is partnering with Data & Society Research Institute to seek a National Forum Grant to create ways for the library community of practitioners, advocates, and scholars to learn, gather, and discuss what the library value of privacy means in the digital world, and to create a roadmap for building national digital strategies that are responsive to patron needs and privacy preferences.

Citizens have become increasingly concerned with their privacy rights online. Due to consumer demand and international pressure, more technology companies and digital services are taking steps to protect privacy, both through policy and design decisions. Not immune to concerns over privacy and protecting sensitive information, librarians and allied information professionals have been equally engaged with policy and technology decisions centered on protecting patron privacy. With a renewed focus on building a national digital platform and critically engage with digital tools, librarians are increasingly confronted with questions on how to consider users' concerns and library values in regards to in-house tools, open source products, search and discovery platforms, and licensed resources.

As libraries and funders build the national digital platform and critically engage with digital tools, how are they considering users' concerns and library values like privacy? How are libraries using patron engagement with our digital presence as moments to showcase digital services and tools that consider users' rights and needs in a more nuanced way?

This project builds upon the recently funded privacy projects and the expertise of the project team from academia, libraries, associations, and nonprofits to convene experts to help answer those questions. The goal of this project is to create ways for the library profession broadly to discuss what the value of privacy means in the digital world and to create a roadmap for how to build our national digital platform, keeping in mind consumer and patron needs and privacy preferences. This project has three parts and several outputs.

- Part one involves developing a series of field guides that clearly lay out important privacy and security issues relevant to libraries developing digital strategies. Possible field guide topics include: privacy by design, internal library information systems, third-party library software systems, cloud-based library systems, public internet and wifi services, licensing of digital content, data security, government information requests, and social media strategies.
- Part two of the project will involve convening the Library Values & Privacy in our National Digital Strategies Summit. This gathering brings experts from inside and outside of the library community to talk about data privacy and security. With field guides in mind, the goal becomes creating a roadmap for librarians, library associations, and funders.
- Part three will focus on writing a report and developing the roadmap, summarizing what was covered at the summit. The Roadmap will clarify how librarians, funders, and technologists can work within the constraints of library values.

The project promises to create a new shared vision of library values in the digital space. The resulting roadmap will help the profession have a clearer path for how to design and develop privacy-centered digital strategies, and impact how we build digital services, how we license materials from third parties, and how the library profession influences their partners in other fields.

Library Values & Privacy in our National Digital Strategies: Field guides, Convenings, and Conversations

1. Statement of National Need

The UW-Milwaukee Center for Information Policy Research is partnering with Data & Society Research Institute to seek a National Forum Grant to create ways for the library community of practitioners, advocates, and scholars to learn, gather, and discuss what the library value of privacy means in the digital world, and to create a roadmap for building national digital strategies that are responsive to patron needs and privacy preferences.

Citizens have become increasingly concerned with their privacy rights online. Due to consumer demand and international pressure, more technology companies and digital services are taking steps to protect privacy, both through policy and design decisions. The desire for more privacy and security played out recently with Apple's decision to protect consumer's privacy by resisting government requests to unlock a suspect's iPhone¹, Google's push to provide end-to-end encryption on its chat and email services², and Microsoft's fight to keep customer data stored outside the U.S. protected³. Not immune to concerns over privacy and protecting sensitive information, librarians and allied information professionals have been equally engaged with policy and technology decisions centered on protecting patron privacy. With a renewed focus on building a national digital platform and critically engage with digital tools, librarians are increasingly confronted with questions on how to consider users' concerns and library values in regards to in-house tools, open source products, search and discovery platforms, and licensed resources. Similarly, libraries are debating how to leverage patron data to provide enhanced and personalized services, without sacrificing patron privacy. Most importantly, in the face of uncertainty about how a new U.S. administration will approach privacy rights online, librarians and their allies are fighting to protect key principles that ensure the protection of online rights.⁴

Now, more than ever, privacy-related projects in librarianship are burgeoning, library conferences increasingly include a focus on privacy within their themes, and many within the library community are engaged in privacy outreach and advocacy. Funded projects like the Library Freedom Project⁵ and the Data Privacy Project⁶ focus on staff training and technical support for libraries. The NISO Consensus Framework for Patron Privacy⁷ brought together librarians, publishers, and library service providers to begin discussions around patron privacy in

¹ <u>https://www.nytimes.com/2016/02/18/technology/apple-timothy-cook-fbi-san-bernardino.html</u>

² https://www.wired.com/2016/05/allo-duo-google-finally-encrypts-conversations-end-end/

³ <u>https://www.bloomberg.com/news/articles/2016-07-14/microsoft-wins-appeal-in-case-over-customers-e-mail-security</u>

⁴ <u>http://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/72207-wary-of-trump-ala-joins-coalition-advocating-key-tech-principles.html</u>

⁵ <u>https://libraryfreedomproject.org/</u>

⁶ <u>http://www.dataprivacyproject.org/</u>

⁷ http://www.niso.org/topics/tl/patron_privacy/

licensed products. ALA and its members celebrate Choose Privacy Week⁸ each year, engaging in policy and practical questions, bringing in speakers and creating guidelines for digital content and initiatives. Researchers have studied licensing language for third party use of patron data and the potential behavior change of patrons and library staff after online privacy and security training.

At the same time, we have teams of people in libraries and in partnership with libraries who are part of building the national digital platform or who are licensing e-content and tools on behalf of libraries. Within these digital areas, we have the capability to consider our changing society and patrons' choices around their online privacy.

While the recent surge in privacy-related activities within the library community is welcome, we see a gap in the conversations we are having about privacy and our digital presence - a knowledge gap, a lack of shared vocabulary, disparate skillsets, and varied understanding. This gap prevents inclusion across the profession and lacks clarity for those responsible for building tools and licensing products. Further, it prevents a complete conversation about how current society's technological and social changes impact our values, our work, our national projects and presence and therefore many of our interactions with those outside of libraries. We need a place and shared resources to help us scope the conversations we need to have about what privacy means to libraries in the digital world. And, the profession needs a roadmap for a shared vision of a digital future that ensures library values.

The goals of this project are directly in line with the goals of IMLS to promote libraries as strong community anchors and trusted entities in their community, and the goals of the National Digital Platform project category to improve digital library tools and impacting investments

2. Project Design

The goal of this project is to create ways for the library profession broadly to discuss what the value of privacy means in the digital world and to create a roadmap for how to build our national digital platform, keeping in mind consumer and patron needs and privacy preferences and the historical values of librarianship. The objectives are to:

- Better understand the areas where the profession is building libraries' digital presence.
- Create shared vocabulary and understanding with emerging technologies so librarians, privacy experts, and stakeholders can discuss our values in light of the social, cultural, and ethical impact of technological developments.
- Engage technology companies, non-profit organizations, and researchers doing privacyrelated work to enhance our understanding of the nuanced issues in digital privacy.

⁸ <u>https://chooseprivacyweek.org/</u>

• Develop a roadmap for librarians, funders, associations, and library partners with a clear pathway for incorporating the values of librarianship such as privacy into the development of the profession's digital presence.

To accomplish these, the project will create field guides and organize a multi--day summit to be held in the spring of 2018 at the Data & Society Research Institute in New York City. These activities will be executed in three phases, a planning and field guide development phase, the multi--day summit, and a reporting and dissemination phase.

Part One: Planning & Field Guides

Part one involves developing a series of field guides that clearly lay out important privacy and security issues relevant to libraries developing digital tools and strategies. These guides will provide an overview of digital tools and environments, outline key implications for the library value of privacy, and provide an environmental scan of what is currently happening across various communities to address these concerns. Possible field guide topics include: privacy by design, internal library information systems, third-party library software systems, cloud-based library systems, public internet and wifi services, licensing of digital content, data security, government information requests, and social media strategies.

During this part of the project, we will host an initial planning meeting at ALA Annual 2016 in Chicago with the project principles and other available partners to map out topics for the field guides. We will begin planning for the Summit taking place in Spring, 2018, including identifying and inviting participants and working with Data & Society on the logistics to host the event in New York City. An additional informal gathering at ALA Midwinter in January 2017 in Denver will provide midstream feedback on field guides and planning for the summit.

Part Two: Convening of the Library Values & Privacy in our National Digital Strategies Summit

Following the release and discussion of the field guides, part two of the project will be convening the Library Values & Privacy in our National Digital Strategies Summit. This one and half day gathering brings together experts from inside and outside of the library community to talk about data privacy and security.

We will invite 20-25 experts who represent different voices and perspectives, drawing from individuals from the following areas:

- Participants from recent grant-funded privacy projects, such as Alison Macrina from the Library Freedom Project, Todd Carpenter from NISO.
- Representatives from ALA's Offices of Information Technology Policy and Intellectual Freedom, such as James LaRue, Marijke Wisser, and Deborah Caldwell-Stone.
- Librarians serving on association committees that focus on privacy, such as Mike Robinson, Helen Adams, and Galen Charlton.
- Technologists working in and around libraries, such as Eric Hellman.
- Technologists and privacy experts outside of libraries, such as Mark Van Hollebeke of Microsoft.
- Nonprofit organizations working on online privacy, such as Vanessa Rhinesmith from Mozilla, and Brenda Strong at the Future of Privacy Forum.

- Open source developers working on library projects.
- Researchers studying the social, cultural, legal, and ethical impact of data-centric technological development, such as danah boyd, Katie Shilton, and Alice Marwick.
- Privacy officers and compliance specialists, such as William Marden from the New York Public Library, and Omer Tene of the International Association of Privacy Professionals.
- Librarians working with patrons and their colleagues on privacy education projects, such as Melissa Morrone and Jennifer Poggiali.
- Library leaders who negotiate licenses or oversee negotiations of e-content and digital tools from third parties.
- Funders of library projects with an interest in including privacy concerns into their current and future agendas, such as IMLS.

We have gathered letters of support (see supplementary materials) from a preliminary set of stakeholder participants. The project team has a wide network, and we have worked on projects with representatives from these categories.

With field guides in mind and experts convened, the goal becomes creating a roadmap for librarians, library associations, and funders. The first part of the summit will allow for short talks, case studies, and provocations to get participants thinking about the impact of data-centric technologies on our society, the values unique to the library profession, actual projects that have considered privacy and data security in their creation, and the role of libraries in the digital world. This will create a conceptual grounding for the interactive discussions in the remainder of day one and most of day two. Topics for interactive sessions will start with the topics covered in the field guides but expand based on current data, privacy, and technology concerns.

During this gathering we hope to spend time asking:

- What is the role of privacy in the library value system today with current and emerging challenges?
- What can be done on the ground as we build out our digital platform to ensure library values are included in the tools we build and the material we license?
- And, how do we use our digital platform to inform our communities about their online privacy?

Part Three: Evaluating, Reporting, and Dissemination

After the summit, the activities for the remainder of the grant period will focus on writing a report and developing the roadmap. The primary audience for these documents will be members of the library community, funders, and associations. The report will summarize what was covered and the discussions from the summit, making the discussions available to a wider audience. The Roadmap will clarify how librarians, funders, and technologists can work within the constraints of library values.

After the summit, we will post a survey to elicit feedback and provide participants with additional opportunities to share information and evaluate the effectiveness of the summit. We will provide summit participants with opportunities to contribute to the final report and the Roadmap by asking for case studies and stories to be included. We will also share early iterations for their feedback.

Outcomes, Assumptions, & Risks

If our project is carried out, we hope to create more informed librarians and stakeholders, who will better understand the technologies used in libraries and their privacy implications. With this knowledge and common vocabulary, a roadmap will exist that helps the profession and its supporters create the future digital presence of libraries that instills the values and ethics of librarianship. Our assumption is that this is an issue that will still resonate with a broader scope of the library profession when the tools and technologies are explained clearly and assists in helping people consider privacy implications directly. While it could be considered a risk, it is clear, given the new political climate, that issues like privacy, access to information, and intellectual freedom will continue to resonate and will grow in interest in the coming years.

Project Plan, Personnel, & Resources

This project will be co-led by Michael Zimmer, Director of the Center for Information Policy Research at the University of Wisconsin-Milwaukee, and Bonnie Tijerina, researcher at Data & Society Research Institute. They will implement and manage the project.

Deborah Caldwell-Stone, the Deputy Director for the American Library Association's Office for Intellectual Freedom will serve as a primary consultant for the project, providing guidance and advice as field guides are developed and serve as reviewers of the field guides. We anticipate adding a member of the library community as a second primary consultant (awaiting approval), and have included funds to support this fourth potential team member.

The team members' expertise is shown below:

- Michael Zimmer is an Associate Professor in the School of Information Studies at the University of Wisconsin-Milwaukee, where he also serves as Director of the Center for Information Policy Research. His research focuses on online privacy, the ethical dimensions of social media, libraries & privacy, and internet research ethics. Zimmer was recently appointed Editor of the ALA's *Journal of Intellectual Freedom and Privacy*.
- Bonnie Tijerina is a researcher at Data & Society, a research institute in NYC. There, she works on projects related to online privacy and ethics in data research. Prior to that, Bonnie worked in libraries for over a decade. She is co-editor of a new LITA Guide, Protecting Patron Privacy, to be published in 2017.
- Deborah Caldwell-Stone is the Deputy Director for the American Library Association's Office for Intellectual Freedom, where she does research, education and professional development projects for ALA's intellectual freedom and privacy programs.

Financial resources will go towards covering travel costs and catering for convenings, travel to ALA for the project leads, salaries to author the field guides, organize the Summit and smaller discussions at ALA Annual, and conduct background research. Other resources include the events planning staff at Data & Society who will be assisting the leads in convening the summit.

Timeline

The project will commence in June 2017, and be completed in one year:

- June, 2017: Grant begins with an informal gathering at ALA Annual (Chicago, IL). Team members will meet with interested members of the library community to brainstorm topics for field guides.
- July, 2017 March, 2018: Field guides produced and released; planning and preparation for Summit
- January, 2018: Discussions and/or presentations at ALA Midwinter (Denver, CO) as well as a small gathering to showcase progress on field guides and brainstorm topics for Libraries & Privacy Summit.
- March, 2018: Libraries & Privacy Summit Gathering (New York, NY)
- March May, 2018: Work on final report and seek feedback
- May 31, 2018: Final report released

Communication, Dissemination, & Sustainability

A project site will exist on the UW-Milwaukee Center for Information Policy Research's <u>website</u>, where the project will be sustained and embedded amongst several initiatives and projects. We plan to publish each field guide on the site and communicate each publication throughout the project through team members' networks, including appropriate privacy-related and technology groups. Reporting and dissemination of this project will involve authoring a report and Roadmap summarizing and distilling the discussion for the broader community. The report and all field guides will also be deposited in the UW-Milwaukee <u>Digital Commons</u>, providing longterm sustainability.

Data & Society has agreed to work with the team and their broader network to disseminate the project's research, guides, roadmap, and discussions. Activities include:

- Bonnie Tijerina and other team members will author posts on <u>Points</u>, Data & Society's Medium site with other 8,400 followers which will be tailored for a broader audience
- The Communications team and project team will gather a series of posts on Points from experts inside and outside of librarianship, meant to inform a broader audience about the work of the team and also as required reading in advance of the summit. These posts will all be clustered and released just before the summit. This "clustering" style has been successfully used at Data & Society with the related of their recent Media, Technology, Politics pieces.
- Data & Society will host informal gatherings during their happy hour for feedback on guides as they are released at Data & Society's Manhattan location.

To ensure the report and findings reach the relevant constituencies, we will be presenting at relevant conferences such as ALA, the Coalition for Networked Information meetings, and the Digital Library Federation Forum. Other dissemination possibilities include publishing summaries of field guides and final report in the ALA's *Journal of Intellectual Freedom and Privacy* (edited by Michael Zimmer), utilizing other official ALA communication channels

(made available via Deborah Caldwell-Stone), and the broader social and professional networks of the various partners and stakeholders involved in the project.

3. National Impact

A national forum is critical at this time because privacy concerns are continuing to grow at the same time that the field of librarianship is trying to figure out their role to meeting needs of users in the digital space while incorporating the values of the profession While many individuals, groups, and institutions are committing resources towards addressing privacy issues in libraries, these efforts are uncoordinated and fragmented. There are conversations happening in different parts of the profession, focusing on discrete topics, but the absence of the overarching conversation about how privacy plays into our digital futures, including how we build in-house tools, utilize open source software, and license tools and services as part of our digital presence is concerning. The profession needs to be having these technical, social, and ethical conversations. This gap prohibits alignment of our values, our work, our national projects in light of online privacy concerns across the profession and prevents clarity for those responsible for building tools and licensing products.

If the goals and objectives of this project are met, there is great potential to create a new shared vision of library values in the digital space. Dissemination of the field guides can support the broader community of practice, helping librarians and information professionals gain a deeper understanding of the privacy issues facing libraries today, connecting them to what is happening outside of libraries in these areas, and guide the development of a national digital strategies. Since we are hoping to support a community of practice with these guides, we will survey those who access the guides and those who attend the summit to measure the effectiveness of the guides.

The roadmap will provide the profession a clearer path for designing and implementing privacycentered digital strategies. Since we will be creating the roadmap based on the conversations and topics at the summit, we will measure success by surveying attendees on the usefulness of the convening as a place to learn from outside experts and focus in on the big topic of what privacy means to the profession in the digital world.

The national impact of *Library Values & Privacy in our National Digital Strategies: Field guides, Convenings, and Conversations* has the potential to be long-lasting and important in this pivotal moment. Building resources and hosting a convening that will make the issues clearer can only help the continual dialogue needed in librarianship to assess our values. The project will impact how we talk about technology and privacy, build digital services, how we license materials from third parties, and how the library profession influences their partners in other fields.



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.

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.

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.

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.

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.

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).

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).

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).

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).

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

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).

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).

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.

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.

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.

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.

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

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

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

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

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.

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

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

Name of publicly accessible source code repository:

URL:

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.

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?

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).

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.

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).

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?

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

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

Name of repository:

URL:

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