

Open-Source Hardware Assembly, Repair, and Sustainability | *University of Washington*

Abstract. The University of Washington Information School proposes a three-year *Laura Bush 21st Century Librarian Early-Career Development* project titled “Open-Source Hardware Assembly, Repair, and Sustainability” that addresses program Goal 2, and Objective 2.3 (per section A.2 of the 2022 NOFO). This \$317,332 grant will support Dr. Nicholas Weber’s investigation of how open-source hardware is being used to meet an increased demand for economic and environmentally responsible computing. The results of this research will positively impact the success of open-source hardware adoption in public institutions and increase the ability of LIS professionals to support open-source hardware as a public interest technology.

Project Justification

Increased access to open-source software and open data have proven to be a positive force for advancing equity - lowering a barrier to educational and research opportunities [19] and providing greater transparency into government actions [37]. LIS researchers and practitioners working in the field of digital curation have been pivotal to the success of both open data and software by creating programming, service models, and documentation that make otherwise complex technologies publicly accessible. IMLS funding has played an important role in these successes - from grants to open-source projects like ePadd, Zotero, and Murkurtu to providing research funds to Lyrasis’ study of open-source software in cultural heritage [1] and digital curation education in civic contexts [2]. This project seeks to understand how LIS researchers and professionals can apply skills in digital curation to increase the success of open-source hardware. Open-source hardware (OSH) is “a device whose design is made publicly available so that anyone can study, modify, distribute, make, and sell the design or hardware based on that design” [3]. Through a set of mixed methods studies, this early-career research proposal will study how documentation impacts the assembly, repair, and sustainability of open-source hardware. In the following sections I further describe open-source hardware and the role that LIS research and practitioners can play in developing devices that meet financial, environmental, and accessibility challenges in contemporary computing.

Open-Source Hardware

In 2005, Nicholas Negroponte introduced the “one laptop per child” (OLPC) initiative with the goal of developing a \$100 device that could connect millions of children to online learning resources [5]. The logic of OLPC was sound: decreasing costs for powerful microprocessors made it possible to both engineer and distribute an ultra-cheap device to students around the world. But ultimately, OLPC failed to produce a reliable device [6]. Many of the initial laptop prototypes were incompatible with sparse energy access in developing countries. And laptops built in the USA used parts that were unavailable in the countries where they were ultimately supposed to be used - preventing meaningful repair of laptops when they broke down [7]. By 2014, the non-profit behind this effort had closed, but the legacy of OLPC - that low-cost, openly accessible hardware could play an important role in computing - lived on in the public imagination.

In 2021, hospitals around the world faced shortages in medical supplies needed to care for Covid-19 patients. Protective equipment, ventilators, and intubation instruments were in high demand, but global supply chains were struggling under the weight of the pandemic in response. Open-source enthusiasts from engineering, biomedicine, and epidemiology - as well as a vibrant maker community - responded with a page from Negroponte’s playbook [8]. The widespread availability of 3D printers made it possible to fabricate PPE at low costs. Sharing open-source CAD files made it possible to both manufacture and repair ventilators needed to keep patients alive [9]. The non-profit Make4Covid delivered 120,000 units of PPE in 6 months, many to rural and children’s hospitals and the Navajo Nation. Open Source Medical Supplies (OSMS), an open-hardware group, created and openly-licensed design plans needed to manufacture ventilators with an estimated commercial value of \$268 million. Volunteers at Helpful Engineering, a medical hardware organization, logged about 23 million hours of volunteer work helping hospitals repair broken equipment, valued at \$130 million in labor costs [10].

The goal of this project is to understand how LIS research and practitioners can help open-source hardware projects, like those produced during Covid-19, avoid the long-term fate of the OLPC initiative. There are a variety of reasons why some open-source projects fail (like OLPC) and others succeed despite long odds (like Make4Covid). One obvious lesson to be learned from the failure of OLPC is that the reliable assembly, repair, and sustainability of open-source hardware depends on high-quality documentation [11]. And yet, accurate part lists, physical design schematics, and software interfaces that are required to reliably build open-source hardware devices are rarely produced [12]. Through a mixed methods research design this project seeks to better understand challenges to creating and maintaining quality open-source hardware documentation. In doing so, I will answer three research questions:

- RQ 1: How do practitioners judge the quality of open-source hardware documentation?
- RQ 2: How does documentation quality impact the success or failure of an open-source hardware project?
- RQ 3: What motivates open-source hardware projects to improve documentation quality?

In the following sections, I further describe open-source hardware and the documentation that is necessary for projects to succeed over the long-term. I then describe how LIS research and professionals can extend existing skills in digital curation to improve the access and long-term sustainability of open-source hardware.

The Source of Open-Source Hardware

Open-source is a licensing scheme for resources that are free to inspect, use, modify, and redistribute. Open-source is also frequently used as a short-hand description of innovative processes for developing and maintaining public goods that rely, in part, on collective action [13]. In both senses of the term, the value proposition of open-source is that reduced transaction costs can lead to rapid innovation, reduce the financial burden of procuring technologies, and create a more equitable distribution of goods for both users and developers. Open-source software (OSS) - like Mozilla Firefox, Linux, Python and LibreOffice - is perhaps the most recognizable digital resource under the open-source label. Critical to the success of OSS is the fact that openly licensed source code can be distributively managed - allowing diverse and often novel contributions from individuals that are decentralized [14].

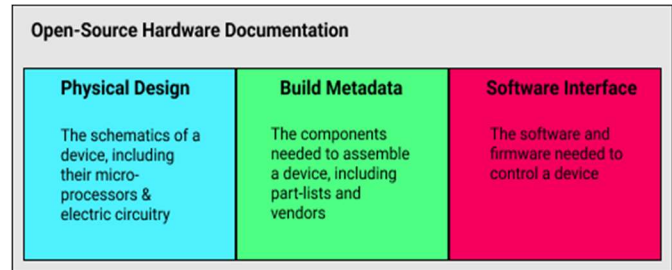
Open-source hardware (OSH) is “a device whose design is made publicly available so that anyone can study, modify, distribute, make, and sell the design or hardware based on that design” [3]. The “source” of open-source hardware is not source code, as in software, but instead **the design and documentation of a device** [4]. Perhaps the most recognizable successes of an emerging open-source hardware ecosystem are *Arduino*, the microcontroller board, and *Raspberry Pi*, a microprocessor-based minicomputer. In the past two years, the number of registered OSH projects has increased 64% [10]. There are a variety of external factors that are contributing to the increased popularity of open-source hardware, including economic, environmental, and legislative challenges for accessible computing.

Open-source hardware offers a viable way to overcome economic challenges: Open-source instrumentation in the field of microscopy has been shown to reduce the costs of basic scientific research by up to 99% [5, 13]. In precision agriculture, low-cost open-source hardware sensors have improved crop yields in Cordoba, Spain by up to 37% [16]. Economically, open-source hardware is not just a cheap alternative to proprietary hardware but can in fact be an economic driver of innovation for the public sector. The EU recently estimated that a public investment of around €1 billion in OSS and OSH in 2018 resulted in an impact on the European economy of between €65 and €95 billion [17].

Open-source hardware can also improve environmental sustainability through legislation: Device manufacturers have increasingly adopted a practice of ‘planned obsolescence’ (the purposeful shortening of device usability) in order to increase consumption and prohibit meaningful repair of hardware. A glut of unrepairable devices creates environmental harms, or e-Waste, that are significant and unevenly distributed - with developed nations often exporting discarded devices to developing countries where chemicals and non-degradable metals impact fresh water sources, and permanently damage ecosystems. In an effort to reduce e-Waste, governments across Europe and numerous US states are currently considering ‘right-to-repair’ laws that would give consumers greater control and protections for repairing hardware devices [18]. While this is good news for anyone who has experienced a faulty laptop battery or a broken phone screen, the implications of these laws for open-source biomedical hardware are also critical for an effective crisis response, as demonstrated by Covid-19, and reducing environmental harms of hardware manufacturing [19].

Open-Source Hardware Documentation: Despite the promise of open-source hardware to solve vexing environmental and economic inequalities in computing, it remains difficult to execute and sustain an open-source hardware project [12]. Whereas open-source software can benefit from shallow fixes (e.g., updating a source library, or fixing a typing error), open-source hardware requires complex design specifications and specific engineering documentation that are essential to a device's success [20]. The three major forms of documentation that are essential for an open-source hardware project to succeed are: **physical design documentation**, **build metadata documentation** (such as parts lists), and **software interface documentation** [21]. Low quality documentation can create a barrier to the adoption and widespread success of open-source hardware projects [22], including challenges in the assembly, repair, and sustainability of hardware. For example - open-source hardware projects are often difficult to reproduce because of missing instructions, necessary part

lists, or lack of specificity in *physical-design documentation* that describes *how* a device is to be built (*Assembly*); The exchange and discovery of relevant open-source hardware is often hurt by a lack of archiving and publishing of OSH *build metadata* that specifies *what* parts can be appropriately used from which vendors [6]. This prohibits non-experts from easily finding and reusing OSH plans as an alternative to proprietary instruments or in repairing existing hardware (*Repair*); and, Many OSH projects require not only timely documentation, but also ongoing work to refresh *build metadata*, and sustain *software documentation*. If these specific pieces of documentation are not updated, then communities are prohibited from engaging in the ongoing development and maintenance of open-source hardware (*Sustainability*) [7].



Extending Digital Curation to Open-Source Hardware

Over the last two decades, LIS researchers and practitioners have developed the sub-field of digital curation to preserve and provide sustainable access to digital artifacts. Digital curation has played an important role in the success of numerous digital initiatives in cultural heritage, civics, and scholarship. One significant example is the role of data curators in helping researchers plan for the long-term management of data [23]. A key activity in ‘upstream’ digital curation is the active and ongoing management of digital artifacts. This includes helping researchers - at the point of data collection - develop and comply with standards for describing data and select existing digital infrastructures to reliably preserve software needed for meaningful reuse [24]. Extending these existing practices to the upstream curation of open-source hardware development has the potential to significantly impact the success of this emerging field by improving the ability to assemble, repair, and sustain hardware devices. LIS research can also help identify substantive barriers to the creation of upstream documentation that improves the potential for reuse of hardware devices in unanticipated contexts. LIS researchers - funded through the IMLS early-career award - have already achieved great success in developing practical approaches to the long-term maintenance of digital heritage systems [25], and the upstream curation of natural history databases [26]. The goal of this early-career research project is to characterize the challenges that open-source hardware projects face in creating high-quality documentation, and to intervene with LIS expertise to help these projects create documentation needed to assemble, repair, and sustain valuable open-hardware devices. The outcomes of this research will inform open-source hardware and LIS practice in assembling and sustaining hardware, as well as LIS educators by creating curricula that helps bridge a gap between hardware knowledge and documentation best-practices.

My long-term research agenda analyzes the success of open-source communities by studying the role of documentation in achieving sustained public-interest outcomes. This research project is a natural extension of my work in digital curation for software and data sustainability. For example, I have previously studied the use of open-data in libraries [27] research [28] and public sector collaborations [29]. In this work, I have described the central role of transparency documentation that fosters cooperation between the public and policymakers [23]. I have studied a variety of open-source communities in previous work [30], but I am also a practitioner in open-source where I lead two large-scale software projects that improve access to public sector data [26, 27]. An extension of this work into open-source hardware will help me lead our field in expanding the impact of digital curation. While the curation of software and data has been critical to the success of digital initiatives, there has been substantially less focus on hardware as a necessary component of achieving success with open infrastructure. This project will help bridge this gap by educating, researching, and applying best practices in upstream digital curation to the development of open-source hardware.

Project Work Plan

In this project, I will complete three studies: in **Study 1**, I will conduct a survey that focuses on understanding challenges in maintaining open-source hardware documentation; In **Study 2**, I will audit the quality of documentation produced by existing open-source hardware projects. By observing the difference between what practitioners report as challenging (survey), and then observing their practices in creating and maintaining documentation (audits) these two studies will produce a holistic picture of challenges in assembling, repairing, and sustaining open-source hardware projects. In **Study 3**, I will execute an experimental intervention to help existing open-source projects improve their documentation practices. Each study builds upon and is informed by the results of the previous. This multi-step research design will deepen insights

on the role that documentation plays in making open-source hardware more accessible to scientists, technologists, and educators. The results of this empirical work will be disseminated in peer-reviewed scholarly publications, as well as LIS practitioner focused reports and auditing frameworks aimed at helping open-source hardware developers improve documentation quality.

Project Research Design

The research design for this project follows the MIDA framework - which requires social scientists to specify information about their background theoretical model (*M*), their motivation for inquiry (*I*), their data collection strategy (*D*), and their analysis strategy (*A*). **Each study, described in-depth below, follows the MIDA framework.** At the outset, it may seem like these studies are complex and too ambitious for a three-year project. But I have already invested significant effort in designing each study with the DeclareDesign platform¹, which provides software for describing, assessing, and conducting empirical research. This pre-research design work enables me to verify that the proposed method of analysis and sample size for each study can produce statistically significant results. Each research study is, in a sense, shovel ready. The methods of data collection and analysis are verified by DeclareDesign, and I have even simulated analysis with synthetic data to ensure that the process of analyzing results will be successful once data is collected.

Before data collection begins, I will also pre-register the research design for each study on the Open Science Framework. Upon completion of each study, all data collected and software produced will be published using an open-access license on Harvard’s Dataverse. Preliminary data collected for survey analysis (Study 1) and for documentation audits (Study 2) has been determined exempt by the University of Washington’s IRB. I will amend this protocol upon the start of this grant to obtain any additional exemption for human subject research data that is collected (Study 3).

	Model	Inquiry	Data	Analysis
Study 1	Documentation Preferences	Survey	Self-reported data (n = ~150)	Common Factor Analysis
Study 2	Documentation Practices	Audit	Observational data (n=390)	Qualitative Comparative Analysis
Study 3	Prosocial Prompt Interventions	Unsupervised Experiment	Observational data (n=390); Self-reported data (n = ~100)	One Way ANOVA

Table 1: An overview of the MIDA framework used to design each research study.

Study 1: Documentation Survey

Model: Uncovering a gap between *what people say* vs. *what they do in practice* is key to applied social science research. And yet, it is all too common for research in open-source to either survey a sample of the community about their practices or conduct observational research that describes what community members do in practice. This research project will attempt to both document what practitioners *say* is important about documentation and then observe hardware documentation *in practice*. The first study in this project will gather information about documentation challenges of open-source hardware practitioners based on a survey questionnaire. Conceptually, there is a good deal of survey evidence that documentation in open-source is neglected because developers often lack time, energy, and understanding of what makes for useful documentation [33]. For example, the social coding platform GitHub annually surveys developers about challenges in using open-source software. For the past three years, over 90% of developers report that incomplete or confusing documentation is the biggest dilemma in using open-source hardware. And yet, 60% of contributors say they rarely or never contribute to documentation [34]. In a recent study about open-source hardware produced by the Wilson Center, experts note that the biggest challenge to increasing capacity and adoption of open-source hardware is a lack of standardization in documenting the design, build, and software needed to make these devices broadly accessible [10].

The goal of my first study is to collect self-reported data from practitioners to better understand challenges they face in creating documentation and the factors they identify as important for high-quality hardware documentation. The research

¹ DeclareDesign <https://declaredesign.org/>

question I will answer through this survey is: ***How do practitioners judge the quality of open-source hardware documentation? (RQ 1)*** Answering this question will provide two important outcomes: 1.) I will characterize hardware documentation challenges based on the perception of open-source practitioners. This will provide a valuable baseline for future work that attempts to intervene in and improve upon the status-quo of documentation practices in open-source hardware. 2.) By documenting what practitioners say is important and then observing their documentation practices in Study 2, I will be able to reliably describe how documentation challenges impact the long-term success of an open-source hardware project.

Inquiry: The Open Source Hardware Association (OSHW), founded in 2012, is a non-profit organization that has played an important role in the professionalization of open-source hardware - from organizing annual research and development meetings to establishing licensing and registration infrastructure for hardware developers. Since 2013, OSHWA has been conducting surveys of practitioners on a variety of topics, including demographic information about who participates in open-source hardware projects, challenges in sourcing parts, and needed educational materials for reliably sustaining open-source hardware devices. Analysis of these surveys has focused on reporting descriptive statistics about open-source hardware communities on a year-by-year basis. In collaboration with OSHWA, I will analyze previous survey data from 5 different community surveys conducted from 2013-2021. As a proof of concept, I have already analyzed demographic information from these surveys to show that, while projects in Africa and South America have increased in each survey, there is still relatively little data on the development and use of open-source hardware in Asia-Pacific regions [35]. Building upon this exploratory demographic analysis, I will design a new survey questionnaire that specifically asks diverse hardware practitioners about challenges they face in creating and documenting open-source projects and how they evaluate the quality of documentation produced by other open-source hardware projects. The survey will be promoted by GOSH², OSHWA, and Open Hardware Makers³ in order to reach diverse international participants that work across Science, Civic, and Educational contexts.

Data: The survey questionnaire will produce individual practitioner responses in the form of a Qualtrics dataframe. The data will be stored as a comma-separated value file that can be analyzed using the open-source statistical platform R-Studio.

Analysis: Survey responses will be analyzed using two methods: 1. Descriptive statistics will be used to describe the central tendency of the responses; and 2. To understand how different variables related to documentation (e.g., frequency of updates, amount, etc.) impact judgements of quality, I will use a common factor analysis. The basic assumption of factor analysis is that, for a 'collection of observed variables, there are a set of *underlying* variables called **factors** (smaller than the observed variables) that can explain the interrelationships among those variables' [36]. Factor analysis, in the context of Study 1, will allow me to describe the relationship between multiple observed variables about what constitutes documentation quality, and ultimately answer: ***How do practitioners judge the quality of open-source hardware documentation? (RQ 1)***

Study 2: Documentation Audits

Model: Previous work has shed light on the critical role of documentation in a variety of open-source communities, including data science [37], software [38] and hardware [39]. These studies make clear that, while general purpose computing can rely on manuals written for a broad audience, an often-overlooked component of sustainable open-source projects is high-quality documentation [40]. Documentation quality in open-source has been evaluated using criteria such as content, structure, and style [41]. In the second study of this project, I seek to better understand documentation in practice by conducting an audit of open-source hardware repositories that are used to store documentation - including physical design, build metadata, and necessary software interface documentation. The goal of this study is to observe, in practice, the way that open-source hardware practitioners create and maintain necessary documentation. The research question I will answer through this audit is: ***How does documentation quality impact the success or failure of an open-***

² Gathering of Open Science Hardware: <https://openhardware.science/>

³ Open Hardware Makers is a non-profit that provides mentoring for new and emerging open-hardware projects: <https://openhardware.space/>

source hardware project? (RQ 2) Answering this question will enable me to make causal claims about the role that documentation plays in successful assembly, repair, and sustainability of hardware.

Inquiry: I refer to the overall approach to evaluating hardware documentation as an ‘audit’ - a technique that is increasingly used to interrogate how complex technologies like algorithms [42], machine learning applications [43] can disproportionately harm particular groups. My goal in this audit is to evaluate the quality of hardware documentation, and the role that documentary practices play in sustaining successful open-source hardware projects. To conduct this audit, I will first construct a sample of open-source hardware projects in three dimensions: The first dimension will focus on the tenure of the project, or how long it has been in existence. I will purposefully select a sample of projects that are less than three years old and projects that are greater than three years old (as measured by the date of a first commit to an open-source hardware repository). The second dimension will focus on the domain application of the hardware. I will purposefully sample hardware projects whose contribution is focused on: Scientific research, Civic Technology, and Education. The third dimension of my sample will focus on a temporal dimension of documentation maintenance. I will purposefully select projects whose documentation has and has not been updated over a 12-month period. This dimension of the sample will shed light on the maintenance of hardware documentation. This purposeful sampling strategy will yield 90 projects that will be audited.

Data: For each project in Study 2’s sample (n = 90), I will audit the available documentation for the physical design, build metadata, and necessary software to assemble an open-source device. This audit will focus on the *content*, *structure*, and *style* of each documentation source using a framework previously developed for open-source documentation quality []. In Table 2 below, I define each dimension of the audit framework. Each dimension of a project’s documentation will be scored on a scale of 1-10, with a protocol that explains the justification for a missing or low-quality documentation score (1) or a complete and high-quality documentation score (10). This process will then be repeated with a group of students in a subsequent round of data collection and analysis (described in detail below in the Directed Research Group section) to expand the total size of the sample to 390 open-source hardware projects.

Content	Structure	Style
Effectiveness: Does the document make effective use of technical vocabulary?	Cohesion: How well does the document connect to other aspects of the device (e.g., software or build metadata)	Appeal: Does the documentation contain information relevant to the intended domain audience?
Consistency: How consistent is the use of terminology?	Conciseness: How succinct is the information provided?	Readability: How accessible is the language of the documentation for the domain audience?
Clarity: Does the document include assembly instructions?	Standardization: Does the documentation reference or use standards?	Understandability: How well does the documentation define or explain concepts?

Table 2: Preliminary audit framework categories for content, structure, and style of hardware documentation.

Study 2 - Directed Research Group: The research activities in Study 2 will also serve as the basis for a directed research group (DRG) at the University of Washington’s Information School in 2023. DRGs at the UW iSchool are for-credit applied learning courses that enable teams of students to work closely with faculty to learn about an emerging technology and apply their learning to an ongoing research project. The DRG will enroll ~15 MLIS students for three quarters of for-credit coursework to participate in two activities: 1.) Students will self-select into groups that will assemble an open-source hardware device in the domain of Science, Civics, or Education (described as part of Study 2’s sample above). Under my direction, each student team will use hardware documentation (physical design, build metadata, and software documentation) to assemble a hardware device: This will include studying a design specification, purchasing necessary hardware components, and then assembling a device. Students will learn skills such as how to locate and purchase hardware components, soldering and electrical circuitry for assembling microboards, and then testing software interfaces to control a hardware device. We will document our experience in assembling the device and report the outcomes of this

work as applied case-studies that describe the role of documentation in building open-source hardware for LIS professionals. 2.) Student teams will also assist in collecting another round of data by auditing the documentation of an additional set of open-source hardware projects (n=300). Students will be trained to use an evaluation framework (described in detail above) for auditing the physical design, build metadata, and software documentation for projects in each domain (Science, Civics, or Education). We will compare our evaluations of this documentation and apply the results of our work to case-study reports that record best-practices for creating and maintaining hardware documentation. **These two activities will play an important role in closing the loop between hardware documentation evaluation and LIS professional skill sets in digital curation that can be assistive to the creation of high-quality open-source documentation.** Under my supervision students will gain hands-on experience assembling open-source hardware and applying their experiences in this activity to the downstream processes necessary to curate open-source hardware documentation for long-term sustainability.

Analysis: After conducting an audit for all 390 projects in Study 2's sample, I will use a method called Qualitative Comparative Analysis to answer the research question: ***How does documentation quality impact the success or failure of an open-source hardware project? (RQ 2).*** Qualitative Comparative Analysis (QCA) is a method developed in comparative sociology and political science that uses Boolean algebra to facilitate logical comparison of case study data that has been coded using subjective scoring criteria [44]. In the context of this project, QCA will enable me to reliably compare open-source hardware projects based on quality dimensions of documentation. This means that, with a sample of 390 projects scored on 10 dimensions, my project will evaluate 3900 data points to determine what aspects of an open-source hardware project correlate with its long-term sustainability, what factors predict long-term sustainability, and how the frequency of documentation maintenance impacts the assembly and repair of hardware devices over time. The sampling strategy I have designed enables these analyses to be conducted based on the type of documentation, the quality of the documentation, as well as the tenure and domain of the open-source hardware project. This study will therefore be able to make causal claims about, for example, what leads to successful outcomes for scientific open-source hardware, and why or how this differs from civic open-source hardware projects. Or even what types of documentation (e.g., build metadata) are hardest to create, and how potential interventions could be designed to improve these documentation practices across domains.

Study 3: An Intervention to Improve Documentation Quality

Model: Experimental behavioral research tests hypotheses by creating purposeful interventions. These interventions can range in complexity from a simple rewording of a message to induce compliance with a policy to manipulating environmental variables that nudge users towards less toxic online speech. Interventions in open-source research have been used to, for example, nudge participants towards more accurate reporting of research results [45] and to improve inclusivity language used in code-of-conduct documents [46] In Study 3, I will design an intervention that is meant to induce open-source hardware projects to improve the quality of their documentation using a 'pro-social prompt.' In social psychology, pro-social prompts are used as a messaging strategy that attempts to persuade an experimental subject to compare themselves to peers before making a decision [47]. So, rather than framing a message in terms of a normative description (e.g., "Your project is bad at physical design documentation"), a prosocial prompt for open-source hardware practitioners will describe peer behaviors (e.g., "Most civic tech projects have software interface documentation in their hardware repository") and attempt to persuade practitioners to improve their documentation as a result. Pro-social prompts have been successfully used in a number of settings, for example to encourage hotel guests to reuse their bath towels based not on what is the right or wrong environmental decision, but instead on what other hotel guests do [47]. In Study 3, I will execute an intervention to test the null hypothesis: *Prosocial prompts will not induce documentation changes based on peer behavior (H₀).* Ultimately, I will use the data from this study to answer the research question: **What motivates open-source hardware projects to improve documentation? (RQ 3)**

Inquiry: The sample of open-source hardware projects in Study 3 will include each project that is audited in Study 2 (n=90) as well as an additional sample of open-source hardware projects audited by the Directed Research Group (n=300). The total sample (n=390) will be split into three treatments in order to test a hypothesis about the effectiveness of prosocial prompts to improve hardware documentation:

- Treatment 1: 130 randomly selected projects will be contacted by email. In this email, I will explain our research into the creation of quality documentation and provide some best practices that have been learned about what aspects of documentation quality are valued by practitioners.

- Treatment 2: A second sample of 130 randomly selected projects will be contacted by email. In this treatment my message will contain a targeted prosocial prompt that both explains our research and provides an analysis of what successful peer open-source projects create in terms of high-quality documentation. Peers will be determined based on the domain application of the open-source hardware project - Science Research, Civic Technology, or Education.
- Treatment 3: The remaining audited projects (n=130) will receive no message. This control group will provide for a baseline to account for projects that may, without prompting, change hardware documentation.

I will then measure the effectiveness of messaging strategies for inducing improvements to an open-source hardware project's documentation. I will do this by using an automated monitoring tool that tracks changes made to a hardware repository [48]. The monitoring tool will not collect any personally identifiable information, but simply observe whether or not public documentation is modified. Monitoring will occur for a 9-month period to account for whether the intervention prompted change, or not. For projects that do change their documentation I will send a short 5 question survey that asks about the effectiveness of the intervention. This short survey will ask projects why they decided to change their documentation, what was difficult or challenging about these changes to documentation, and the overall impact that the practitioner believes documentation plays in the success of their hardware device.

Data: The two treatments described above will produce a dataset containing metadata about each project (e.g., the domain application, the tenure of the project, etc.) as well as the effect of a treatment. For projects that do change their documentation, my survey questionnaire will produce a Qualtrics dataframe that can be stored as a comma-separated value document for analysis.

Analysis: To analyze the resulting data from the two treatment effects, and ultimately reject or fail to reject Study 3's hypothesis, I will use a one-way ANOVA test for significance. ANOVA is used to assess statistical significance of 'differences among observed treatment means based on whether their variance is larger than expected because of random variation' [49]. In Study 3, we are testing the null hypothesis that prosocial prompts will not induce documentation changes based on peer behavior. By observing the difference in the three mean treatment groups I can then reliably reject or fail to reject the null hypothesis. The combination of two sources of evidence will enable me to reliably determine the effectiveness of the intervention, as well as capture motivations for documentation change. This study will then answer **RQ 3: What motivates open-source hardware projects to improve documentation?**

Project Staff

Funding for full-time post-baccalaureate researchers will allow me to significantly improve the quality of research that can be conducted in a short three-year window. With full-time research assistants, I will be able to collect and analyze data across multiple studies, and design and execute interventions that improve the success of open-source hardware projects. One of my primary career goals is to increase access to higher education for students that might not otherwise have the opportunity. I have practiced this as the director of a Public Interest Technology clinic at UW that provides summer internships for students across the Puget Sound Region. This research project would substantially further this goal by providing funding for two post-baccalaureate fellows that are purposefully recruited from undergraduate minority-serving institutions (MSIs). The post-baccalaureate fellows will each be trained and mentored in conducting research into documentation and open-source communities over a 1.5 year period. In total, this funding will prepare two fellows (over three years) for advanced graduate research in an LIS PhD program, with the goal of each student continuing at UW as my PhD advisee.

Advisory Board

To ensure that my project is meeting the needs of libraries, educators, and open-source hardware experts I have convened an advisory board consisting of diverse experts (see attached letters of support). Each advisory board member will be paid \$500 per year to attend two annual teleconferences. During these conferences, I will provide project updates and solicit feedback on project deliverables. Advisory board members will produce a short evaluation of the project's progress each year. The advisory board will also provide mentorship on project design and assure that research and education opportunities are in-service of both hardware and LIS practitioners. The board will consist of:

Julieta Arancio - Researcher with OpenFlexure Microscope at University of Bath (UK) and Postdoc at Drexel University. Dr. Arancio is an expert in OSH development, and the lead organizer of the Open Hardware Makers project.

Alicia Gibb - Executive Director of the Open Source Hardware Association (OSHW). Ms. Gibb holds an MLIS degree from Pratt and is an expert on the licensing and production of open-source hardware.

Jenny Molloy - Founder and Director of the Bioeconomy Lab at Cambridge; Board Member of the Gathering of Open Science Hardware (GOSH). Dr. Molloy is an expert in scientific instrumentation and the economics of open hardware.

Maggie Melo - Assistant Professor at University of North Carolina's School of Information and Library Science. Dr. Melo is an expert in makerspaces for libraries. She is also an innovative educator that has developed curriculum for fabrication, hardware, and making in LIS degree programs.

Diversity Plan

Diversity has been considered in terms of participants and project sampling, advisory board membership, and the creation of professional training materials that can be used by both open-source hardware and LIS practitioners.

Participant diversity is central to understanding how and in what ways open-source hardware increases access, lowers costs, and levels educational barriers that depend upon computing. This project uses purposeful sampling in the selection of projects that will be audited in Study 2. My sampling strategy includes selecting geographically diverse projects from Africa, Asia, and the Global South, as well as sampling project documentation across Scientific, Civic, and Educational domain applications. These projects collectively serve diverse audiences and increase the equity of computing access. Improving the likelihood of success for these projects has a cascading impact on the likelihood of open-source hardware adoption.

In the education sector, access to hardware is often a significant barrier to participation in STEM training [50]. By investigating, documenting, and easing the repair of open-source hardware this project has the potential to significantly improve the diversity of basic research practitioners in STEM. In designing pro-social prompts for improving the timeliness of open-source hardware documentation this project will improve the potential for existing open-source hardware devices to be used in colleges, universities, and places of learning throughout the US. Further, I have purposefully selected open-source hardware projects that serve diverse audiences - including projects like the open-hardware microscopy which provides design templates for building devices used in community college biology courses. The results of Study 3 will provide LIS researchers an increased capacity to not only design impactful interventions, but also equip practitioners with concrete guidance on how to assist hardware developers in creating useful documentation for the maintenance of an open-source project over time.

The results of Study 1 and Study 2 will produce an open-access report on best-practices in open-source hardware documentation. The report will be shared with and promoted by hardware professional societies and advocacy groups, such as OSHWA, GOSH, and Open Hardware Makers. Based on the results of Study 2, I will also work with Open Hardware Makers - a non-profit organization modeled after the Mozilla Open Leaders program - to create a self-auditing protocol that can be used to help new open-source hardware projects evaluate the quality of their documentation. This outcome has the potential to improve not just how open-source hardware projects build documentation, but also positively impact the development of new devices, and strengthen the uptake of these devices for students and researchers working in under-resourced contexts.

In the fall of 2023, I will lead a directed research group (DRG) focused on open-source hardware. This DRG will recruit diverse students through the UW McNair Scholars program. The DRG will engage students in the assembly and auditing of supporting documentation for open-source hardware devices. Students will gain knowledge about how to build hardware and the role that documentation plays in the process of sustaining an open-source project. These practical skills will create a cohort of MLIS students that are capable of immediately impacting the field's contributions to hardware sustainability and increased access to open-source technologies. Curriculum used to teach students basic skills in soldering, wiring, and software interfaces will be shared on a project website and will be licensed for free and open reuse by other LIS instructors, enabling the field of LIS to begin meaningfully engaging students in hardware education.

Project Results

Outcomes	Explanation	Beneficiaries
Best-Practices Report	An open-access report that summarizes the findings of all three studies, and concisely explains the implications of documentation quality on open-source hardware	OSH practitioners, LIS practitioners
Self-Audit Framework	A framework based on Study 2's outcomes will be described in a short report, and provided for OSH practitioners to use in evaluating and auditing their documentation	OSH practitioners, LIS practitioners
Research Publications	Study results will be reported in three separate papers submitted for peer-review at LIS journals, and presented at LIS professional conferences	LIS researchers, LIS practitioners and educators
Case Studies of OSH Assembly in LIS	Case studies will be published as open-access documents on the project website	OSH practitioners, LIS practitioners, researchers and practitioners
Curriculum	Recorded lectures, class activities, and assembly exercises for open-source hardware will be published to the project website	LIS practitioners, LIS educators

Table 3: Overview of project results

To sustain the outcomes of this project I will deposit all data, software, and publications in open-access repositories. I will publish a project webpage that will serve as a hub for communicating about the project with beneficiaries, provide a project roadmap, share preliminary findings, and recruit potential collaborators. This project will conduct three studies of the role that documentation plays in assembling, repairing, and sustaining open-source hardware devices: In Study 1, a survey of documentation practices, will collect and analyze survey data about the challenges that open-source hardware practitioners report in creating physical design, build metadata, and OSH software documentation. **The results from this work will shed light on the perceptions of documentation by OSH practitioners and provide a needed baseline for understanding how LIS skills in digital curation can be usefully applied to the emerging practices in open-source hardware.** These results will be summarized in an open-access report that is disseminated to open-source hardware communities and LIS practitioners, and summarized in a research article for peer review in the open-access Journal of Open Hardware.

In Study 2, an audit of documentation practices, will evaluate the quality of hardware documentation. This audit will provide results about not just what practitioners say, but how they document an open-source hardware device. **The results of analyzing audit data will allow me to report the correlations between domains, project tenure, and documentation maintenance on the success or failure of an open-source hardware project.** These results will inform a self-audit framework that is provided to open-source hardware practitioners through the Open Hardware Makers organization. A research paper summarizing the findings will be deposited in the open-access preprint server ArXiv and submitted for peer review at the Journal of the Association for Information Science and Technology (JASIST).

In Study 3, a behavioral intervention will use prosocial prompts to encourage open-source hardware practitioners to improve documentation related to a hardware device. **The results of this work will provide data about motivations for documentation improvements, as well as an intervention which attempts to induce 390 open-source hardware projects to improve their documentation.** A paper summarizing the effectiveness of this intervention will be submitted for peer review in the open-access journal PLoSOne.

Extending digital curation skills in LIS to the domain of hardware can significantly improve the ability of freely accessible devices to be used in education, research, and civic contexts. The results of this project can help IMLS achieve goals towards diversifying the field of practitioners and researchers by providing educational materials to train future practitioners in curating hardware documentation. The research results of this work will also impact equity by improving documentation quality for existing open-source hardware projects - enabling these devices to be both more accessible and useful to civic, education, and research practitioners.

Digital Products Plan

Type

What digital products will you create?

Comma-separated value (data) - Data collected from surveys, audits, and an intervention will all be stored as UTF-8 encoded CSV files. CSV is an open format that is small in size, and can be used across operating systems and applications. The size of this digital project will be ~200mb. CSV files will each be described using the Project Open Data Metadata schema 1.1. The size of this digital project will be ~200mb

Qualtrics survey questionnaire (project documentation) - Two surveys will be designed and hosted on the Qualtrics platform. At the conclusion of data collection on Qualtrics we will export a digital copy in both PDF and JSON to archive. Both file formats will be described using the Project Open Data Metadata schema 1.1. The size of this digital project will be ~50mb

Research design (project documentation) - This project is committed to transparency throughout its execution. I have designed each research study using the DeclareDesign platform that allows for a protocol and registration template to be generated based on the methods of data collection, sample size, and intended methods of analysis. Research design documentation will take the form of PDF documents. The research design documents will be described using the Open Data Metadata schema 1.1. The size of this digital project will be ~50mb

Software - This project will create R-based Jupyter notebooks in order to analyze data collected in surveys, audits, and an experimental intervention. R is open-source programming language that is widely used in statistical analysis, and the Jupyter notebook ecosystem allows for literate programming - where software code, narrative explanation, and data are combined to increase transparency of research. Software developed for this project will be described using the CodeMeta vocabulary v2.0. The size of this digital project will be ~2gb.

Project website - A project website developed using the React framework, and will be published using Github Pages. All source files will be version controlled using a public Github repository. The site will be archived at the conclusion of the project. Site metadata will be recorded in a YAML document. The size of this digital project will be ~2gb.

Availability

How will you make your digital products openly available (as appropriate)?

Data - CSVs of data collected during Study 1, 2, and 3, project documentation, and instruments developed (e.g. survey questionnaires) will be initially stored in a password protected server at UW, and backed up to an encrypted cloud infrastructure. Upon completing data collection for each study, we will remove any potential identifying information about participants and archive the data using a CC-BY license in a Harvard Dataverse collection that is dedicated to this project. These data will be free to reuse and openly accessible to the public.

Software - R-based notebooks will be stored and version controlled in a public Github repository under an MIT license. Software will be freely available and openly accessible to the public throughout the entirety of the grant. At the conclusion of the grant we will deposit Software in a Harvard Dataverse collection that is dedicated to this project.

Website - the project website will be published openly on the web, and website source code will be stored and version controlled in a Github Repository under an MIT license. The source code will be openly accessible to the public and free to reuse.

Project design - In order to conform with best practices in open-science my research project will generate a protocol for collecting, analyzing and reporting research results. These PDF documents will be submitted and archived as pre-registrations on the OSF platform. OSF pre-registration submissions follow the DataCite metadata standard. The pre-registrations will be assigned a CC-BY and openly accessible to the public and free to reuse.

Access

What rights will you assert over your digital products, and what limitations, if any, will you place on their use? Will your products implicate privacy concerns or cultural sensitivities, and if so, how will you address them?

Data + Project Documentation - Data and documentation generated by this project will be anonymized and then published to a Dataverse collection under a CC-BY license that requires acknowledgement of reuse. There will be no other rights asserted over the reuse of data from this project.

Software + Website - R-based notebooks, scripts, and the project website will have an MIT license. MIT grants all rights to use, study, change, and share source code.

Sustainability

How will you address the sustainability of your digital products?

Data + Project Documentation + Software - Upon completion of the project, anonymized data, documentation, and software will be archived at a Harvard Dataverse collection dedicated to this project. Harvard's Dataverse is an open-access free repository that is certified by the Core Trust Seal foundation. Data will be preserved indefinitely, and freely accessible upon completion of the grant.

Data Management Plan

Identify the type(s) and estimated amount of data you plan to collect or generate, and the purpose or intended use(s) to which you expect them to be put. Describe the method(s) you will use, the proposed scope and scale, and the approximate dates or intervals at which you will collect or generate data.

This project will collect self-reported data from surveys as well as administrative data that is collected by evaluating public documentation available on the web. These two data sources will be analyzed for reporting descriptive statistics. The data collected will approximate 200 mb of storage that will be securely managed using UW and cloud infrastructure. All data will be stored on a password protected server managed by the University of Washington's Information School. A copy of all data will be backed up using an encrypted drive that is stored on Dropbox. Research design information, including software and documentation used in analysis will be stored on a password protected server managed by the University of Washington's Information School, and version controlled in a public repository on Github. Data will be collected and stored in the comma-separated-values (CSV) format, with plain-text UTF-8 encoding. Data collection will begin in August of 2022 and end in March of 2025.

Will you collect any sensitive information? This may include personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information. If so, detail the specific steps you will take to protect the information while you prepare it for public release (e.g., anonymizing individual identifiers, data aggregation). If the data will not be released publicly, explain why the data cannot be shared due to the protection of privacy, confidentiality, security, intellectual property, and other rights or requirements.

Data resulting from surveys may contain some personally identifiable information, such as IP addresses of participants, or institutional affiliations. We will not use this information in reporting results from research, and will remove this information before archiving data at the end of this project (described in detail below). PII will be removed manually by the Project Director before deposit, and this work will be reviewed by data curators employed by the University of Washington Library.

What technical (hardware and/or software) requirements or dependencies would be necessary for understanding retrieving, displaying, processing, or otherwise reusing the data? How can these tools be accessed (e.g., open-source and freely available, commercially available, available from your research team)?

This project will produce data stored in CSV formats that are accessible using any desktop computing environment with a text editor. The software that we will use to analyze this data will be in the R programming language. All scripts, analytic procedures, and data processing that are produced in R will be created in reproducible R-based Jupyter notebooks, and stored in a version controlled repository on Github.

What documentation (e.g., consent agreements, data documentation, codebooks, metadata, and analytical and procedural information) will you capture or create along with the data? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the data it describes to enable future reuse?

Data will be described in structured metadata using Project Open Data Metadata schema v1.1 (<https://resources.data.gov/resources/dcat-us/>). Variables in quantitative data collected during Study 1, 2, and 3 will be described using a data dictionary that is formatted as CSV. A codebook will be developed to analyze data from Study 1 and 2. The codebook will be stored as a CSV, and will be described with a data dictionary (and in narrative that accompanies our design protocol described below). Software will be described in structured metadata using the CodeMeta schema. Each of the three studies in this project have been designed using the DeclareDesign platform (<https://declaredesign.org/>). This tool provides a research design template, synthetic data that models the method and sample size reported in the study, and a set of R scripts for analyzing collected data. These research design documents will be included in a pre-registration of our study on Open Science Framework, and stored in a version controlled repository on Github. In advance of collecting or analyzing data, the protocol and design for all three studies will be pre-registered on the Open Science Framework (OSF <https://osf.io/prereg/>)

Metadata for data and software, research design templates, and pre-registration documents will be stored in a public version controlled repository on Github, and backed up on server at the University of Washington. At the conclusion of

the project all metadata, data and software will be deposited at Harvard University's Dataverse. Each metadata and documentation file will be assigned a Digital Object Identifier and linked to the overall project repository.

What is your plan for managing, disseminating, and preserving data after the completion of the award-funded project? If relevant, identify the repository where you will deposit your data. When and for how long will data be made available to other users?

Data from all three studies will be managed by the Project Director (Weber). He will be responsible for all documentation, storage, preservation, and archiving of data. Upon completion of the project in July of 2025, all data, software, and accompanying documentation will be deposited and published in Harvard University's open-access repository Dataverse - <https://dataverse.harvard.edu/>. The data will be assigned a CC-BY license to enable open-access and reuse with attribution. Software will be assigned an MIT license which allows for the open access, modification, and distribution of the software. Data and Software produced by this project will remain accessible in perpetuity.

When and how frequently will you review your Data Management Plan? How will the implementation be monitored?

The data management plan for this project will be reviewed twice a year in advance of an advisory board meeting. Any updates to the data management plan will be documented and version controlled in a public software repository on Github that is accessible to anyone with an internet connection. Updates to the data management plan will be reported to the advisory board as well as in an interim project report submitted to IMLS and published to the public repository on Github.

Organizational Profile

The Information School at the University of Washington

Mission

Our Passion. We are inspired by information. We want everyone to know how vital information is in all aspects of life.

Our Vision. We envision a world where effective use of information helps everyone discover, learn, innovate, solve problems and have fun. We envision a world free of existential problems. Information changes lives. *Our Mission.* We make information work. We prepare information leaders. We research the problems and opportunities of information. We design solutions to information challenges.

- Source: <https://ischool.uw.edu/about/mission-vision>, Adopted by the Dean most recently in 2021 after consultation with faculty, staff, student and external advisory boards and councils.

Governance Structure

The Information School is one of 18 independent schools and colleges comprising the University of Washington, a Tier 1 public research university ranked by Reuters as one of the top five most innovative public universities in the world in 2019. Study at the iSchool is guided by the Dean, who reports to the Provost. The iSchool currently consists of 70 faculty members of diverse expertise, with backgrounds ranging from the library and computer sciences to education, business, philosophy, and sociology.

Service Area

The UW iSchool serves the people of the state of Washington.

Brief History

Founded in 1911, the library school at the University of Washington was established as a response to the growing need, in the Western United States, for highly trained, well-prepared librarians. Over the course of the next 90 years, the school continued to play an essential role in the field of librarianship in the Northwest, as the school gained a reputation for producing extremely strong library professionals. Beginning in 2000, in response to changes in the ways people create, store, find, manipulate and share information, the school introduced a variety of new continuing education certificate programs and new degree programs, including the Online Master of Library and Information Science, the Bachelor of Science in Informatics, the Ph.D. in Information Science and the Master of Science in Information Management. In 2001, the Information School became the newest independent school of the UW, known simply as the Information School, or the iSchool, for short.

In the most recent U.S. News and World Report rankings (2021) of Library and Information Science programs, the UW iSchool is ranked second overall in the nation; second for digital librarianship and for information systems; third in health librarianship; fifth in services for children and youth; and eleventh in school library media. As a leading member of the iSchool movement, the UW is a model for other information schools around the world. The iSchool offers four degree-granting programs. The flagship program, the ALA-accredited Master of Science in Library and Information Science (MLIS), the oldest such program west of the Mississippi River. The iSchool also offers a Bachelor of Science in Informatics, Master of Science in Information Management (MSIM), and PhD in Information Science.